Fewer Licenses, Similar Teachers:
Changing Licensing Tests in Indiana

John Fallon*[1] and Marcus A. Winters, Ph.D.†[1]

[1]Boston University

March 11, 2025

Abstract

We use longitudinal administrative data from Indiana to examine changes in teacher quality following the state's shift to a more stringent licensure test. Despite a significant drop in new licenses issued following the change in the licensure test standard, the overall quality of incoming teachers and the relative quality of licensed teachers compared to unlicensed teachers remained largely unchanged. We find some heterogeneity by subject and school setting, with urban schools experiencing a modest decline in teacher quality, particularly in math. Our findings raise questions about the value of requiring prospective teachers to pass licensure tests to obtain a license.

*fallonj@bu.edu
†marcusw@bu.edu

Introduction

In 2014, Indiana replaced its longstanding licensure test for prospective public school teachers with a new test designed to better align with the state's content standards. Policymakers did not intend the new test to be more difficult, but in practice, the change substantially raised the barrier to obtaining a teaching license. Nonetheless, analyzing longitudinal administrative data on the universe of Indiana public school students and teachers, we find little practical difference in the quality distribution of teachers hired under the previous standard and the more stringent new standard.

Every state mandates that public school teachers pass at least one test to obtain their initial license or endorsement for teaching in specific areas. By passing these tests, teachers theoretically demonstrate minimal competency in pedagogy, classroom management, and subject matter knowledge. Licensure standards vary substantially across and within states over time (Larsen et al., 2020; Angrist and Guryan, 2004). The expected impact of changing licensure test requirements on average teacher quality is theoretically ambiguous. While more stringent licensure requirements might increase average teacher quality by eliminating the least qualified prospective teachers, the heightened barriers to entry could simultaneously repel the most promising candidates by diminishing the profession's appeal relative to other career paths, thus reducing average teacher quality (Larsen et al., 2020). Lawmakers have in the past both raised and lowered licensing standards in the name of boosting teacher quality and, thus, student achievement (Larsen et al., 2020).

Ours is the first study to directly measure differences in measurable teacher quality following a change in licensure test standards using a statewide administrative dataset, which allows us to address key limitations of prior studies in this literature. Previous analyses of administrative data find significant but fairly weak correlations between licensure scores and value-added but do not compare differences across multiple licensure regimes (Cowan et al., 2020; Goldhaber, 2007; Shuls and Trivitt, 2015; Shuls, 2018; Clotfelter et al., 2006, 2007a,b, 2010; Strauss and Sawyer, 1986). Prior studies that leverage within-state variation in licensure stringency find no significant impact on the average academic preparedness of entering teachers but do find improvements at the bottom tail of the distribution (Angrist and Guryan, 2004; Larsen et al., 2020). However, these studies lack a direct measure of teacher quality and thus must rely on limited proxies, such as the competitiveness of colleges that produce new teachers.[1]

Our study is most closely related to Chung and Zou (2022)'s recent evaluation of the impact of statewide adoption of edTPA on student outcomes. Leveraging across-state variation in the timing of adopting the more rigorous licensure standard within student-level data for representative samples on the NAEP test, the authors find that edTPA substantially reduced the number of graduates from a state's teacher preparation programs and had small negative impacts on average teacher quality. We build within this literature by providing a detailed assessment of changes in the quality distribution of teachers hired under different licensure test standards experienced across a state and by measuring these differences at the tails of the teacher quality distribution.

Another strength of our paper is that concerns about the quality of the teacher pipeline did not bring about the policy change we study. While some documents discussing the new tests specified that the new tests should be rigorous, the primary reason given for the policy was to align better with content standards adopted in previous years. At minimum, this suggests

[1] Hanushek and Pace (1995) present a cross-sectional analysis of state-level aggregate data and report that college students in states with higher certification standards are less likely to enter teaching.

that the timing of the policy shift is plausibly exogenous, if not the policy change as a whole. Even if one takes issue with the assumptions necessary to believe our estimates are causal, our paper provides descriptive evidence that is difficult to reconcile with the idea that increasing the stringency of licensing tests would improve incoming teacher effectiveness.

We use longitudinal administrative data on public school teachers and students to measure differences in the quality and composition of Indiana public school teachers who gained their initial license under substantially different testing requirements. The number of content areas covered by new licenses granted by the state dropped by about 40% in the first year of the change. Public schools across the state increased their reliance on emergency or provisionally licensed teachers. Despite such substantial differences in the characteristics of the entering teacher workforce, we observe at best marginal differences in the distribution of teacher quality for those who entered under the more stringent licensure test standard. Teachers hired under the more stringent test standard had slightly lower average value-added (-0.007σ) than those hired under the previous licensure regime, and the relative quality of licensed and unlicensed teachers did not differ significantly across the testing standards. However, we find some evidence that the lack of difference in average teacher quality masks a modest increase in the left tail of the quality distribution among English/Language Arts (ELA) teachers and a modest decline in the right tail of the quality distribution of math teachers under the more stringent standard.

We find some evidence that the teacher quality distribution changed somewhat differently following the change in licensure tests in urban public schools than in other settings across the state. Relative to teachers hired under the previous standard, the average value-added for teachers hired under the more stringent test requirement was modestly lower (-0.026σ) in urban schools and modestly higher (0.009σ) in suburban schools. The pattern of results we identify is generally consistent with Larsen et al. (2020)'s model, which predicts that the effect of raising the licensure standard on the top end of the quality distribution is declining in the district's real wage.[2]

We also contribute novel estimates for quality differences between certified and uncertified teachers across two licensure test regimes. We find that certified teachers are, on average, more effective than uncertified teachers under both testing standards. However, after adopting the more stringent standard, the relative benefit of having a certified teacher declined in rural schools but increased in suburban and urban schools. Public school districts have long relied on provisional and emergency certifications to fill open teaching positions when they cannot successfully assign an appropriately licensed teacher, and this practice expanded considerably since the COVID-19 pandemic (Slay et al., 2020; DeArmond et al., 2023). Prior research suggests that conventionally certified teachers have similar or marginally better value-added than emergency or uncertified teachers (Bacher-Hicks et al., 2023; Olivia L. Chi, 2024; Backes et al., 2024) though, at least in some settings, teachers who enter through alternative certification routes (e.g., Teach for America) are as effective or more effective than conventionally certified teachers (Backes and Hansen, 2018; Penner, 2021; Xu et al., 2011; Henry et al., 2014; Clark and Isenberg, 2020; Master et al., 2023).

Our estimates are consistent with several studies finding that licensure requirements within a variety of occupations have null or modest impacts on average labor quality (Carroll and Gaston, 1981; Kleiner and Kudrle, 2000; Kugler and Sauer, 2005; Hall et al., 2019; Kleiner and

[2] The higher "wage" subsumes differences in earnings and non-monetary differences such as quality of the working environment. Under Larsen et al. (2020)'s model, it is ambiguous whether the effect on the bottom of the quality distribution is higher or lower when real wages are higher.

Soltas, 2019; Farronato et al., 2020). Anderson et al. (2020), which serves as a notable recent exception, found that the licensure of midwives in the early 20th century reduced maternal mortality. Further, our findings are consistent with evidence that licensure requirements can raise the floor of workforce quality even if they do not impact the mean (Ramseyer and Rasmusen, 2015; Bhattacharya et al., 2019; Larsen et al., 2020).

Context

Our analysis focuses on the testing component of Indiana's licensure requirements.[3] Indiana began requiring prospective teachers to pass a proficiency test prior to licensure in the mid-1980s. Teachers must pass the pedagogy tests associated with the grade level they will teach. Several teaching positions also require prospects to pass a content area test associated with their subject area or specialization (e.g., special education). Teachers only need to pass these tests once, but if they want to add a new content area license, they must pass the associated test. Teacher candidates pay a fee to take the test, which is currently $114 per test administration.

From the inception of licensure testing until 2014, Indiana used tests created by the Educational Testing Service (ETS), now known as Praxis II. The Praxis II battery of tests assesses knowledge of specific subjects and general and subject-specific teaching skills. Prospective teachers typically take these tests during the final year of their preparation program. Half of

[3]Indiana uses a three-tiered licensure system progressing from Initial Practitioner to Proficient Practitioner and then to Accomplished Practitioner. The Initial Practitioner license is valid for two years and is renewable up to two times. To be eligible for an Initial license, applicants must hold a bachelor's degree, complete an approved teacher preparation program, and pass state-required tests associated with their teaching position. Alternative pathways are available to those with bachelor's degrees but not complete a teacher preparation program. Those pursuing alternative pathways must eventually meet all standard licensure requirements but can fulfill them while teaching in a classroom setting.

all states currently include Praxis II as part of their licensure requirements, and there is little variation in the passing thresholds applied across states.

The Indiana legislature adopted new educator standards in 2010. The state contracted with Pearson to construct 61 new pedagogy and content-area tests that better aligned with the state's new teaching standards. The new test requirement went into effect on February 10, 2014. The state granted prospective teachers a grace period from February to June 10, 2014, allowing them to take either Pearson or Praxis II tests, after which new licenses were granted only based on passing the Pearson test.

First-time passage rates were much lower than policymakers had anticipated.[4] Indiana adjusted the passing thresholds in 2015, but the failure rate remained higher than under Praxis II. Media coverage at the time described considerable frustration with the difficulty of the Pearson test among prospective teachers and administrators and complaints that the requirement was contributing to persistent teacher shortages across the state. In September 2021, Indiana reverted back to using the Praxis II tests for teacher licensure.[5]

Failing to obtain a license does not prohibit an individual from teaching within a public school. As in many states, since the 1960s Indiana has issued emergency teaching permits to address teacher shortages. Teachers on an emergency permit must hold a bachelor's degree from a regionally accredited university and commit to work towards achieving a license. Schools can renew emergency permits if the prospective teacher still meets the requirements. In addition, schools often assign teachers who are not appropriately licensed to teach in areas where they cannot identify an available appropriately licensed teacher. Nationwide, during the 2020-21

[4]For example, see (Segall, 2017)

[5]Ideally, we would also examine the switch back to the Praxis II tests, but we would have very few years of observation, and it would be concurrent with a COVID-19 shock.

school year, about 2% of public school teachers were uncertified, and about 5.5% were teaching on an emergency or provisional certification.[6]

Data

We analyze longitudinal administrative data from the Indiana Department of Education containing information for the universe of students and teachers within the state. A unique classroom identifier matches students to their teachers. Data on students include demographics for each year from 2012 to 2019, classifications to receive supplemental services (e.g., special education, English language learners), and scores on statewide standardized math and ELA tests, which we standardize by grade and year.[7] We match teacher employment records to data documenting all new teaching licenses and renewals after 2012. Because we cannot confidently assign licensure status for those with active licenses before 2012, our estimation samples include only teachers we first observe teaching in 2013 or later.[8]

We do not directly observe the test that a teacher passed to gain their license; thus, we rely on the timing of the license to classify whether the teacher would have been required to have passed the Praxis II or Pearson test. We classify teachers who received a license before February 2014 as licensed under the Praxis regime and those first licensed later as licensed

[6] Digest of Education Statistics 2022, Table 209.26

[7] The state test was the ISTEP from 2012-2018 and the ILEARN in 2019.

[8] We do correct for the possibility that a teacher earned a license prior to 2012 , but started teaching in 2013 or later. If a license starts before 2012 but is active in 2014, we cannot observe that license directly. Licenses can last up to ten years, potentially impacting all years we observe. We try to account for these hidden licenses using the action cited when creating a new license. If a teacher renews a five-year license, we assume the teacher held a license in that area for the five years before renewal. If a teacher is professionalizing a license that requires the teacher to have two years of experience, we assume the teacher held a license for two years before professionalizing. In addition, we restrict our analysis to teachers we observe starting teaching in 2013 or later. For us to not observe this teacher's license, the teacher would have to have earned their license multiple years before beginning teaching.

under Pearson. We classify teachers into one of three categories according to their licensure status in the first year we observed them as teachers. A "properly licensed" teacher holds a license that covers all of the content areas required to teach the class and grade they are teaching; a teacher who is "licensed out of subject" has at least one teaching license of some kind, but not the content area needed to teach the grade and subject for their course;[9] an "unlicensed teacher" includes those with an emergency permit,[10] a substitute permit, a non-teaching license, or no license/permit.

We restrict the data to one observation per subject for each student in each year. To identify the teacher responsible for a student's test score, we rank the various math and ELA classes by frequency for each grade in each year. For each student, we keep the observation for the most frequently taken class. We exclude students who have multiple teachers for the same class.[11]

Table 1 reports descriptive statistics for students and teachers employed in our analyses. About 41.9% of observations are taught by a teacher who was licensed after the adoption of the Pearson Core tests. Test scores were standardized to the Indiana student population prior to restricting the sample, but similar characteristics still hold with the mean test score close to zero and the standard deviation close to one. The vast majority of the teachers are appropriately licensed for the class they are teaching (89.9%), with unlicensed teachers representing 8.3% of observations.

[9] Specifically, a teacher is designated as licensed for a class if they hold an active license in the areas specified by the Indiana assignment codes (Indiana Department of Education, 2023).

[10] Licenses that last one year or less are assumed to be as emergency permits

[11] While we believe that this method is the most correct, computationally feasible method of assigning the teacher responsible for a student's test score. In table 2, we show that randomly assigning a student to one of the teachers with whom they take a class does not change the takeaway from the analysis.

Differences in the Composition of Licensed Teachers

There was a sharp reduction in the number and scope of new teaching licenses across the state following the change from requiring teachers to pass Praxis II to requiring them to pass the more stringent Pearson battery of licensure tests. Figure 1 illustrates the number of content areas [12] covered in new licenses the state granted annually from 2012 through 2019. The total number of content areas on new licenses granted by the state dropped about 47%, from 14,292 during the last year of the Praxis II test to 7,637, for the first entering cohort required to pass the Pearson tests. The decline occurred across all license types and was especially prevalent for licenses to teach ELA.

The transition to the Pearson licensure tests also coincided with substantial changes in the composition of licensed and unlicensed teachers instructing students. Figure 2 describes the annual licensure composition of the teaching workforce for first-year teachers statewide and by residential location, respectively. We observe a notable increase in the proportion of teachers who lack an appropriate permanent license, which began in 2015 and continues throughout the sample period. The proportion of unlicensed teachers appears to have increased in each residential setting.

The increase in emergency licensed teachers appears to have sufficiently addressed any additional shortage of certified teachers that followed the transition to Pearson. Figures C.1-C.5 in the Online Appendix show that the overall pupil-teacher ratio remained steady statewide and within urban, suburban, and rural school systems.

Finally, Table 3 compares the descriptive characteristics of teachers hired and licensed under each testing regime for the years we observe. Relative to the final two years of the Praxis II

[12] A teaching license can cover multiple subjects and grade ranges, each one we define as a "content area."

regime, teachers who entered under the Pearson testing standard were about 1.3 years younger on average. The age difference is specific to licensed teachers.[13] Teachers who began under the Pearson standard are also significantly less likely to be white and more likely to be Black. This change in the racial composition for entering teachers appears to be a product of schools' increased reliance on unlicensed teachers, a higher proportion of whom are Black.

Differences in Average Teacher Effectiveness

Overall Teacher Quality

We first investigate the average difference in student test scores when instructed by a teacher who obtained their initial license under the Praxis II or Pearson testing standards. The estimation samples include observations of students instructed by any teacher who was first hired in 2013 or later. Our initial analyses consider differences in teacher quality overall without distinguishing their entering licensure status.

We begin with a regression comparing average test scores for students instructed by teachers hired under the Praxis II or Pearson regimes. Formally, we estimate the following equation:

$$Y_{icgsjt} = \alpha + \beta PEARSON_j + \chi X_{icsjt} + \gamma f(Y_{icsjt-1}) + \lambda_c + \psi_g + \varepsilon_{icgsjt} \qquad (1)$$

Where Y_{icgsjt} is a student test score in either math or ELA for student i associated in subject c within grade g of school s taught by teacher j during year t; $PEARSON_j$ is an indicator for if a teacher entered under the Pearson test regime; X_{icsjt} is a vector of observed student, school,

[13]Results reported in Figure 3 suggest that an increase in the likelihood that teachers are hired within the first year of receiving their initial license likely drives the drop in age of new teachers under Pearson.

and classroom characteristics, including an indicator for the teacher's years of prior experience. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language learner, or special education designation change relative to the previous year. The function $f(Y_{icsjt-1})$ is a cubic polynomial of the z-score of the test score in both math and ELA separately interacted with the subject of the class as well as dummy variables for missing a test score in a previous year; λ and ψ are fixed effects, respectively, for subject and grade; and ε is a stochastic term. We are primarily interested in β, which measures the conditional difference in average test scores for students instructed by a teacher who entered under the more stringent Pearson testing standard relative to a teacher who entered under the Praxis II standard. [14]

We cannot prove that if Indiana had continued to use Praxis II tests teacher effectiveness would have remained the same. But it is unlikely that there was a sizable improvement in teacher quality as a result of the switch to a more stringent set of tests. If teacher effectiveness would have remained similar in a world where Indiana did not switch to Pearson tests, [15] then switching to the Pearson test caused the observed change (or lack of a change) in effectiveness. For the switch to the Pearson test to have caused an improvement in teacher effectiveness, our analyses would have to be significantly biased downwards.

For our analyses to be biased against the Pearson tests improving teacher quality, counterfactual

[14] We do not control for certain teacher demographics such as age or race. If teacher effectiveness is impacted by changes to the teacher pool, then that is part of the overall impact on students. If anything, Figure 4 shows that the youth of new teachers under the Pearson test regime biases the effect slightly upward as the difference in effectiveness is more negative with more experienced teachers.

[15] Assuming everything else stayed the same.

teacher quality would have to decline around the policy shift, but the evidence does not support this story. We are unaware of any contemporaneous policy changes that would substantially alter the composition of newly licensed teachers. We estimate the effect by the cohort of teachers licensed each quarter. The estimated effectiveness of teachers before the policy shift is stable and shows no evidence of a decline in teacher quality (Figure 5). Our sample only compares teacher effectiveness after the policy shift to about two years' worth of newly licensed teachers. We expand our sample to include much older teachers for whom we are less certain about when they first earned their teaching license. We impute their latest first year of teaching from their experience levels. In this sample, teacher effectiveness is relatively stable, but there is evidence of a gradual improvement in teacher effectiveness before the change to the Pearson tests (Figure 6). This trend suggests that our analyses may be slightly biased in favor of finding an improvement in teacher effectiveness rather than being biased against a positive outcome.

Table 4 reports the results from estimating Equation 1 for the full sample and within subsamples of interest. Overall, the substantial changes in the licensure status of entering teachers described in the previous section did not coincide with a meaningful change in average quality for new entrants. On average, receiving instruction from a teacher who entered under the Pearson regime is associated with a statistically insignificant and immaterial (-0.007σ) drop in student test scores.

We find some evidence of heterogeneity across students and school types. We observe significant but small reductions in average student test scores associated with receiving instruction by a teacher licensed under the Praxis regime for non-white students (-0.019σ), students eligible for free or reduced-priced lunch (-0.012σ), and students enrolled in urban public schools (-0.026σ). We find precisely estimated null differences related to the standard under which the teacher was initially licensed for white students.

Differences Among and Between Licensed and Unlicensed Teachers

Though we do not observe differences in average teacher effectiveness overall, it is possible that the new tests altered the quality of teachers who gained licensure relative to those observed teaching on an emergency license. Such analyses speak more directly to the extent to which the tests differ in their ability to link licensure to teacher quality. Further, altering the relative quality of properly-licensed and emergency-licensed teachers would be especially important from a policy perspective if the state were to limit schools' ability to fill positions with emergency-licensed teachers.

Thus, we consider how much the average effectiveness of licensed and unlicensed teachers differed across cohorts licensed under Praxis II and Pearson test standards.

We add to Equation 1 interactions between license status and testing regime. The regression distinguishes between teachers who at the time of initial hire were properly licensed, unlicensed, or had some but not all necessary licenses to teach within their current area.[16] Formally, we estimate a regression taking the form:

$$Y_{icsty} = \alpha + \sum_{l}^{3} \phi_l LIC_{icsty} + \sum_{l}^{3} \beta_l PEARSON_{ty} * LIC_{icsty} + \chi X_{icsty} + \lambda_c + \psi_g + \varepsilon_{icsty} \quad (2)$$

Figure 7 illustrates the overall effect and 95% confidence interval for each license-status and

[16] We separate licensed teachers teaching out of subject despite typically having too small samples to produce precise estimates because it is not clear whether such teachers would be better incorporated into the unlicensed or licensed categories. While they have demonstrated the ability to pass a licensure test, they have also not demonstrated the capacity to teach within their current area under the licensure regime.

test-regime combination relative to properly licensed teachers who entered under the Praxis II standard (the omitted comparison group). Looking first at the results from the full sample, we observe no significant difference in the relative impacts of licensed and unlicensed teachers across the state. Under the Pearson standard, outcomes for properly licensed teachers declined slightly (-0.005σ), and the average outcomes for students instructed by unlicensed teachers did not change significantly. Under both standards, students statewide score about 0.022σ higher when assigned to a licensed teacher relative to when they are instructed by an unlicensed teacher.

However, the pattern of results again differs across residential settings. Within urban schools we find insignificant and modest declines in the average effectiveness of both unlicensed (-0.028σ) and properly licensed (-0.017σ) teachers. Within suburban schools the average effect of having an unlicensed teacher declined under the Pearson standard by about -0.044σ, but the average effect of licensed teachers remained unchanged. Meanwhile, rural schools experienced a 0.028σ improvement in the relative average quality of unlicensed teachers but no difference in the average effectiveness of properly licensed teachers. For both suburban and rural schools, we observe no difference in the relative quality of unlicensed and licensed teachers across the testing regimes, but the difference in average impacts of unlicensed and licensed teachers increased within urban schools.

Differences at Tails of Teacher Quality Distribution

We now investigate whether Indiana's change in licensure test standard altered the tails of the quality distribution for newly hired teachers. To do this, we calculate a teacher-subject-year level value-added measure. See Appendix B for a detailed description of our approach to estimating teacher value-added, which is similar to the current standard in the literature.

Figure 8 compares the unconditional value-added distribution for all teachers, initially-licensed and initially-unlicensed teachers hired under the Praxis II and Pearson regimes. The lines on the figure illustrate the 10th, 50th, and 90th percentile across the two distributions. Overall, there is not significant movement in the quality distribution for entering teachers. In the cities, there is a decrease in the quality of incoming teachers, which is the largest at the median. In the suburbs, there is a narrowing of the distribution with shrinkage at both the upper and lower tails.

The lack of movement in the combined teacher quality distribution masks some modest changes within each subject. Figure 9 shows that the right tail among math teachers shifts left while the left tail among ELA teachers shifts right.

Figure 10 compares the value-added distributions among licensed and unlicensed teachers across the testing regimes. Relative to the Praxis II standard, the variation in value-added among licensed teachers appears to have decreased slightly under Pearson as the left tail shifts inward. The variance of value-added among unlicensed teachers also shrank and the median shifted slightly upward.

Table 5 allows for inference when comparing different points on the illustrated value-added distributions by reporting the results from conditional quantile regressions testing differences across testing standards at the value-added distribution's 10th, 50th, and 90th percentiles. While these regressions account for experience and subject composition in ways that Figures 8 -10 cannot, the results are stylistically similar.

The limited action we observe across licensure standards occurs in predictable ways at the tails of the teacher quality distribution. Relative to the Praxis II standard, under Pearson among licensed teachers the bottom 10th percentile of value-added increased by about 0.027σ, and the 90th percentile decreased by about -0.012σ. We find a similar decline at the 90th percentile and increase at the 90th percentile for all teachers, although neither is large enough to be statistically significant. This general pattern of results holds if we restrict the sample to only those we have observed teaching for more than two years, giving some confidence that substantial changes in attrition patterns for those licensed by different regimes do not primarily drive the result.

Among teachers in city schools, the quality of the median math teacher dropped by 0.056σ, while the ELA value-added did not change significantly at the 10th, 50th, or 90th percentiles. Among teachers at suburban schools, the 90th Percentile value-added for math teachers declined by 0.063σ, while the 10th percentile of ELA value-added improved by 0.053σ. In rural schools, math value-added declined to a non-significant degree at each percentile, while value-added for ELA improved by 0.032σ at the median.

Conclusion

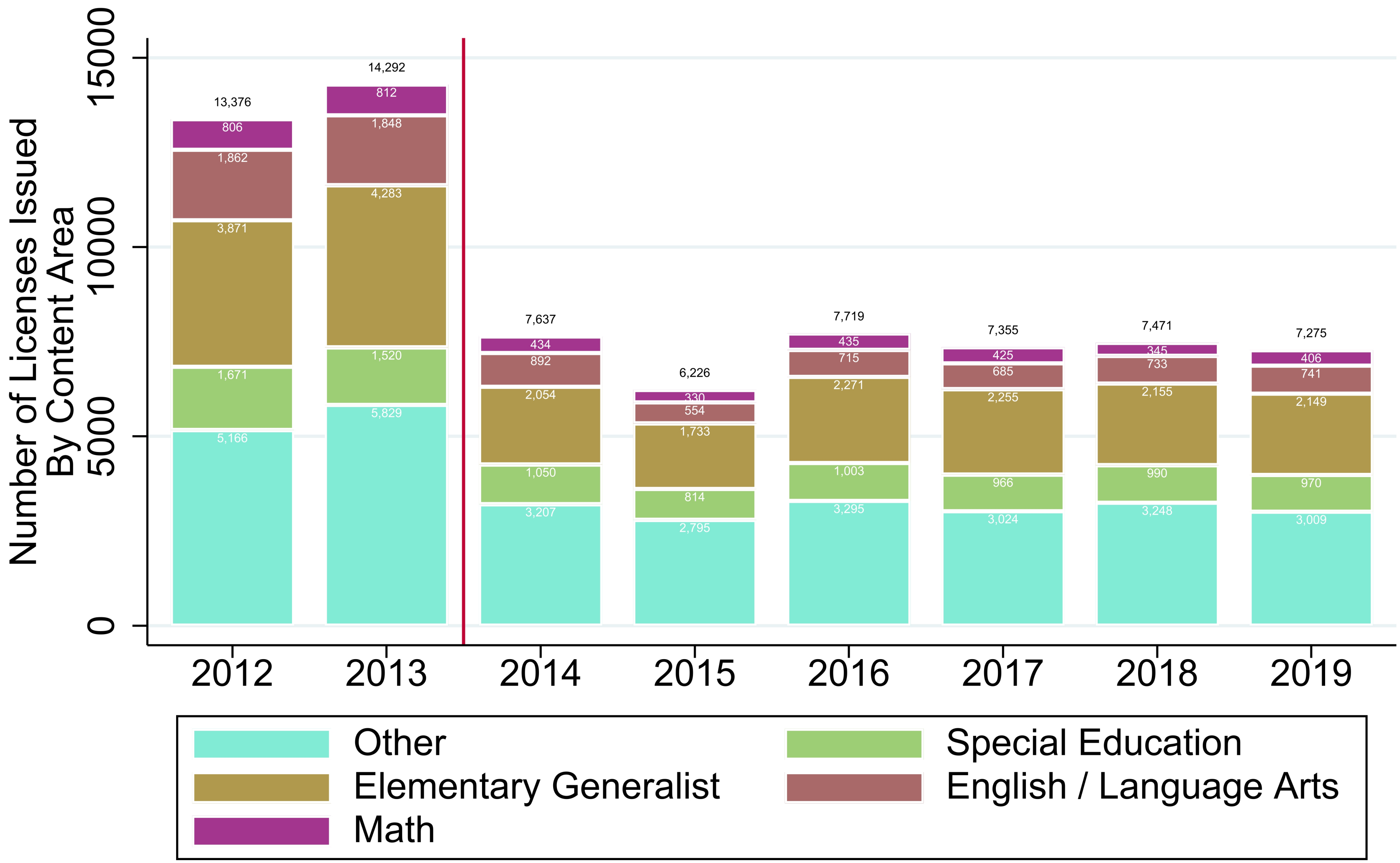
In this paper, we measure differences in the characteristics and quality of entering Indiana public school teachers hired under meaningfully different licensure test standards. After switching to a more stringent standard, teachers hired in Indiana were substantially less likely to have a proper initial license. Despite such changes in the licensure composition of entering teachers, we find little to no difference in the overall value-added distribution statewide. Even within teachers who start licensed, we do not find a meaningful shift in effectiveness.

However, we find some intriguing evidence that changes in teacher quality under the two licensure standards differed somewhat by type of residential location and by classroom subject. Our results suggest that the value-added of incoming licensed and unlicensed teachers tended to decline for urban public schools, while it remained unchanged or increased slightly for suburban and rural schools. Within subject, the quality distribution of math teachers experienced a decline in the top tail of the quality distribution while the quality distribution of English teachers saw an increase in teacher quality at the bottom tail. Though the differences were modest, our results suggest that to the extent that teacher quality changed after strengthening the licensure test standard it tended to disadvantage urban public schools and declined more in math than English.

We contribute to a broader literature on teacher licensure unique evidence derived from longitudinal administrative data in which we are able to compare teachers who entered under two different licensure test standards on a direct measure of teacher quality. Our findings are generally consistent with previous cross-sectional studies that correlate teacher value-added with scores on licensure tests (Cowan et al., 2020; Goldhaber, 2007; Shuls and Trivitt, 2015; Shuls, 2018; Clotfelter et al., 2006, 2007a,b, 2010; Strauss and Sawyer, 1986), studies that leverage variation in licensure test standards over time but use measures of teacher preparation as an imperfect proxy for teacher quality (Angrist and Guryan, 2004; Larsen et al., 2020), and Chung and Zou (2022)'s recent findings for edTPA derived from longitudinal student-level data from a nationally representative sample. Considered together, this evidence calls into the question the value of licensure testing for recruiting a highly effective education workforce.

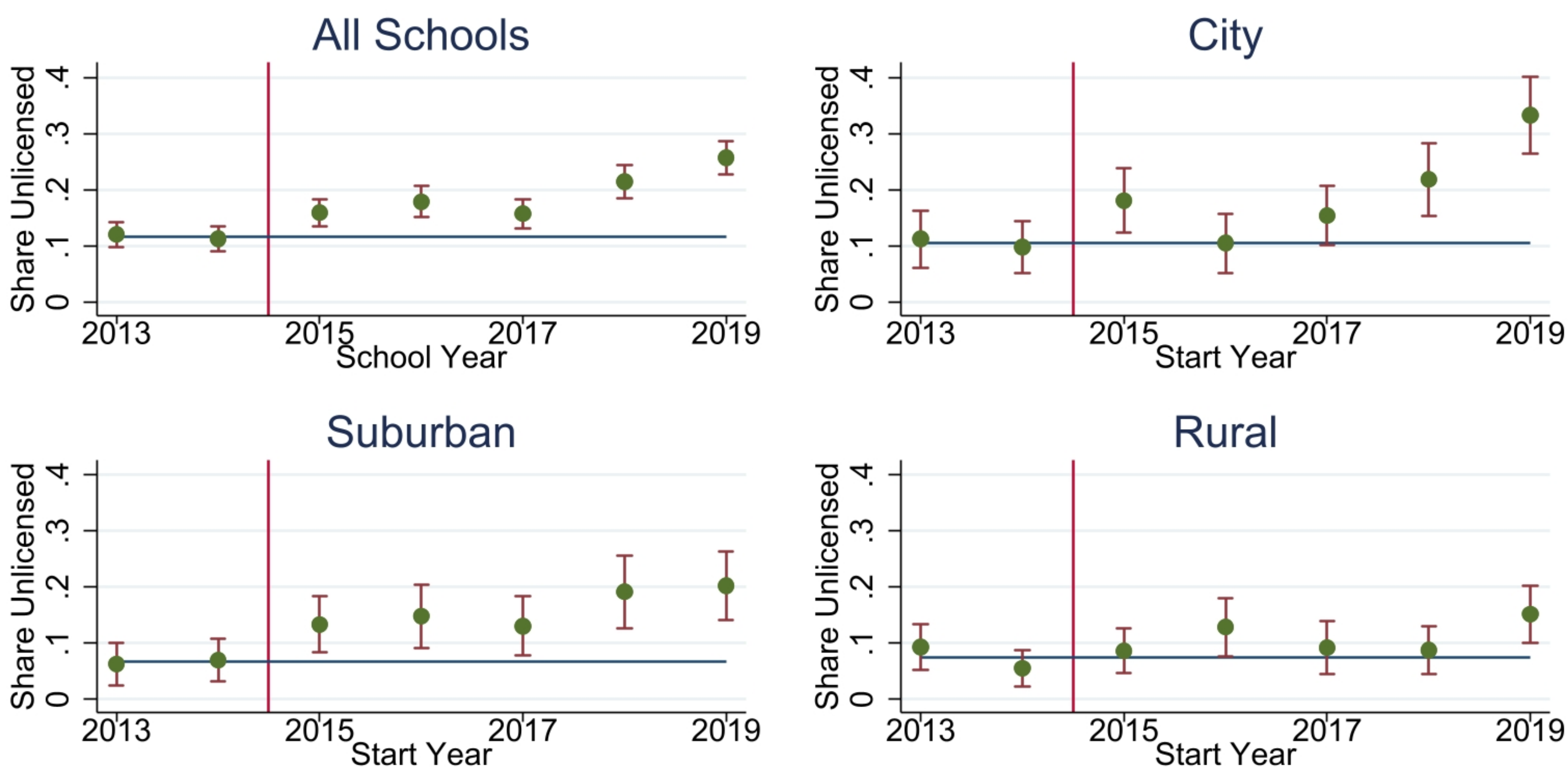
Nonetheless, we caution that we analyze only the effect of changing the state's licensure testing standard, and testing is only one important component of teacher licensure in Indiana and elsewhere. Other licensure requirements may have different impacts on the composition and quality of the entering teacher workforce. For example, Larsen et al. (2020) finds some evidence that increasing the stringency of course requirements leads to more desirable teacher candidates, while changing the stringency of licensure tests does not. However, evidence for the extent to which pre-service training requirements contribute to the quality of the teaching workforce is currently limited. We encourage similar studies that use longitudinal administrative data with direct measures of teacher quality and observations under different licensure standards for pre-service training to examine the extent to which non-test components of teacher licensure impact the supply of public school teachers.

Figure 1: Count of Content Areas Covered By New Licenses By Year



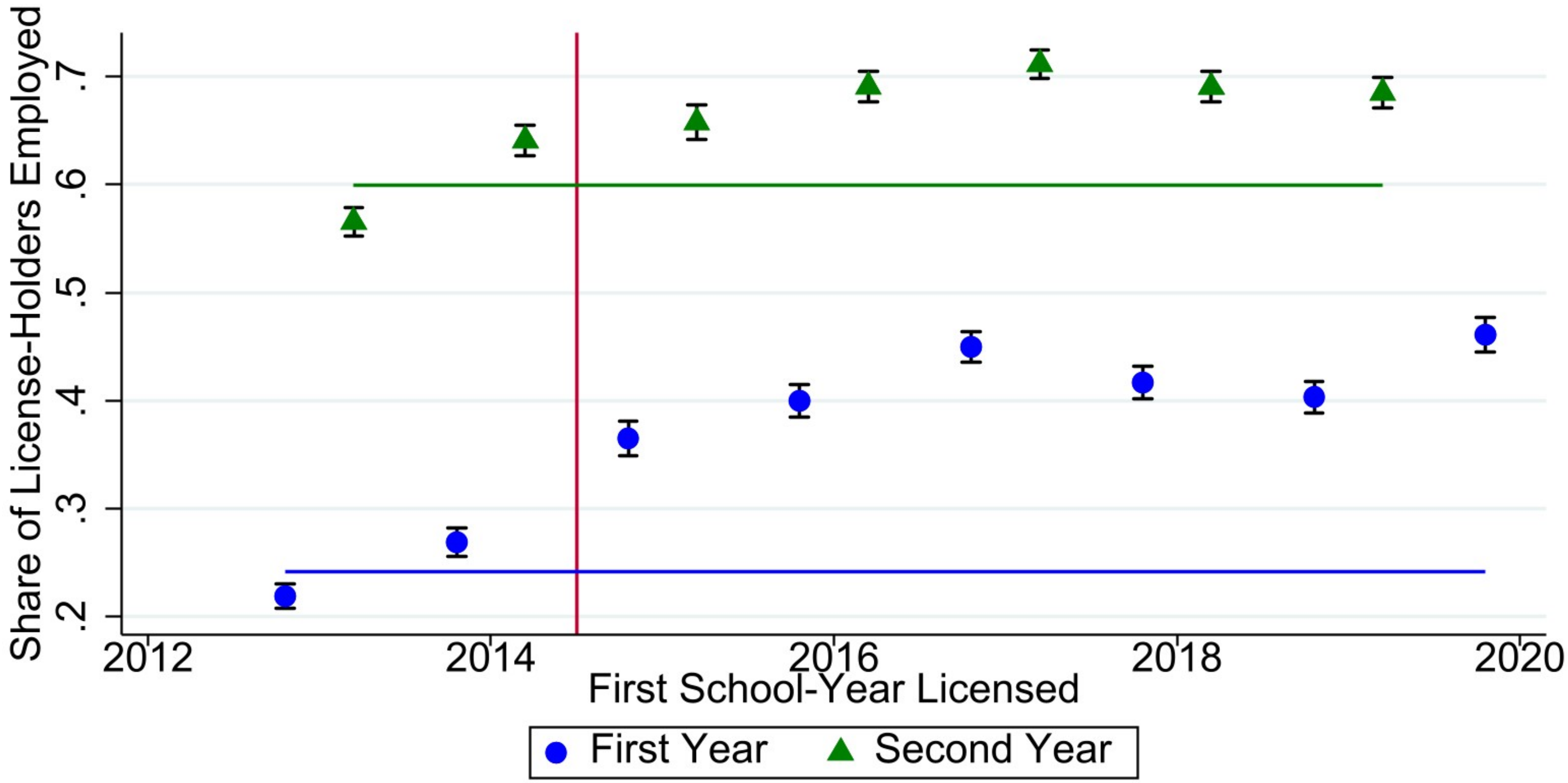
Counts for 2012 are generated using partial years. These counts are extrapolated by dividing by the average share of annual licenses in the available months. We calculate the average share using calendar years 2013-2019. Each content area is categorized exclusively with math and ELA content areas categorized first, followed by science, special education, elementary generalist, and finally other.

Figure 2: Share of Novice Teachers Without a License Over Time



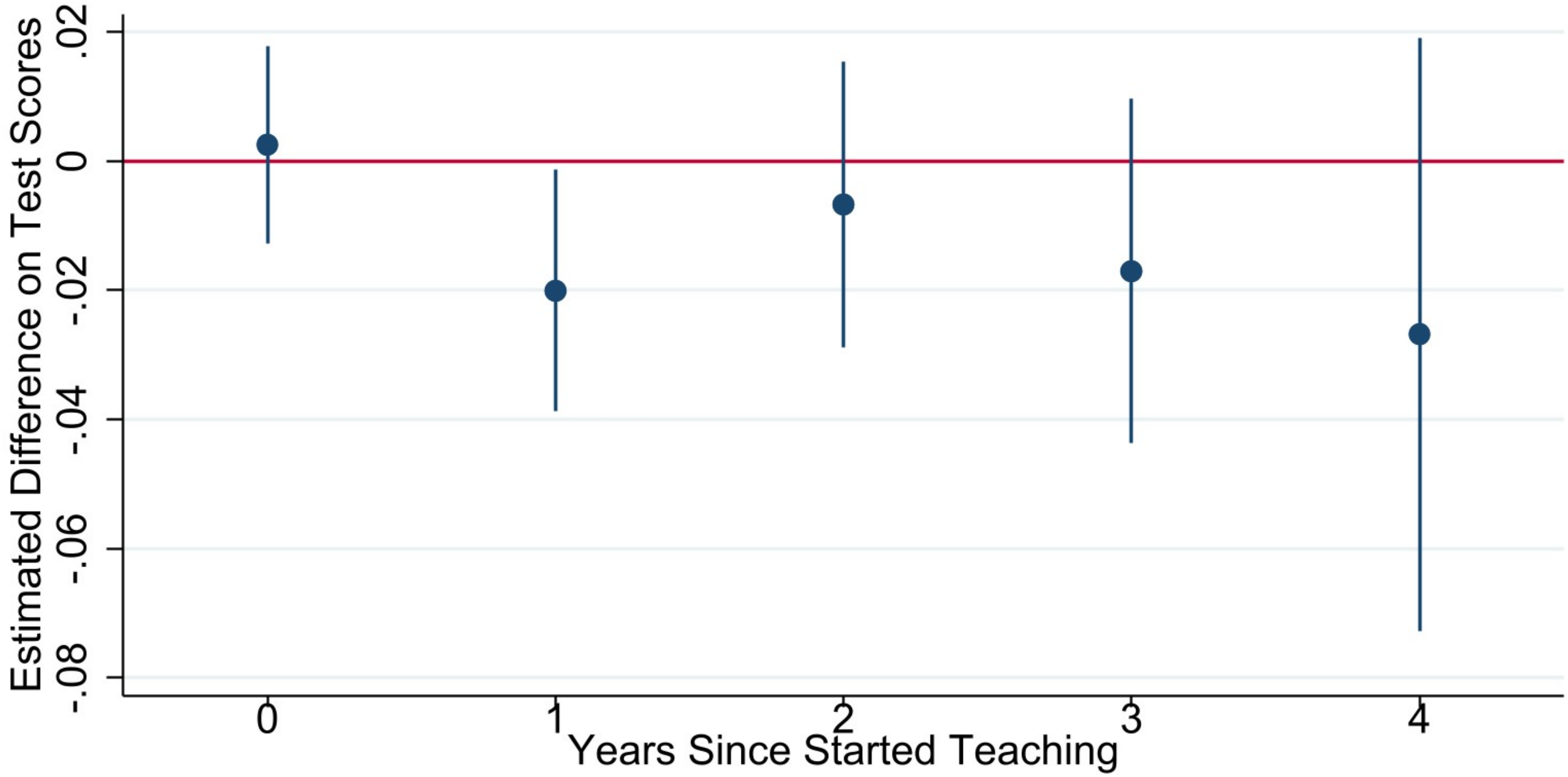
Each point represents the share of new unlicensed teachers each year. All teachers who started teaching after 2012 and are still teaching in the school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.

Figure 3: Share of Licensed Teachers Hired in First Two Years by Year First Licensed



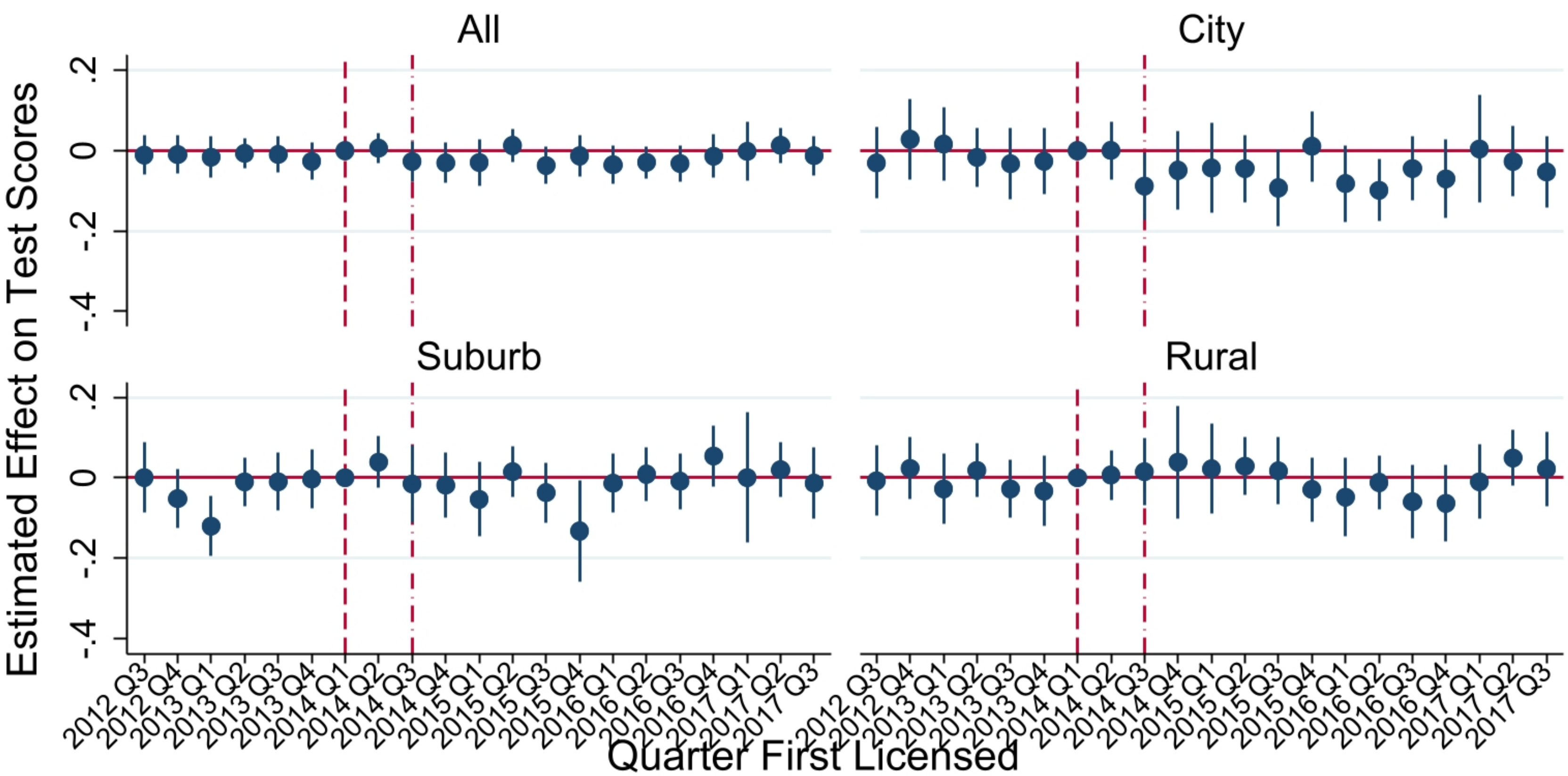
First School-Year Licensed is the first school year a teacher would hold a valid license. If a teacher leaves teaching after starting, we do not consider them employed.

Figure 4: Difference in Student Test Scores of Pearson Licensed Teachers By Years of Experience



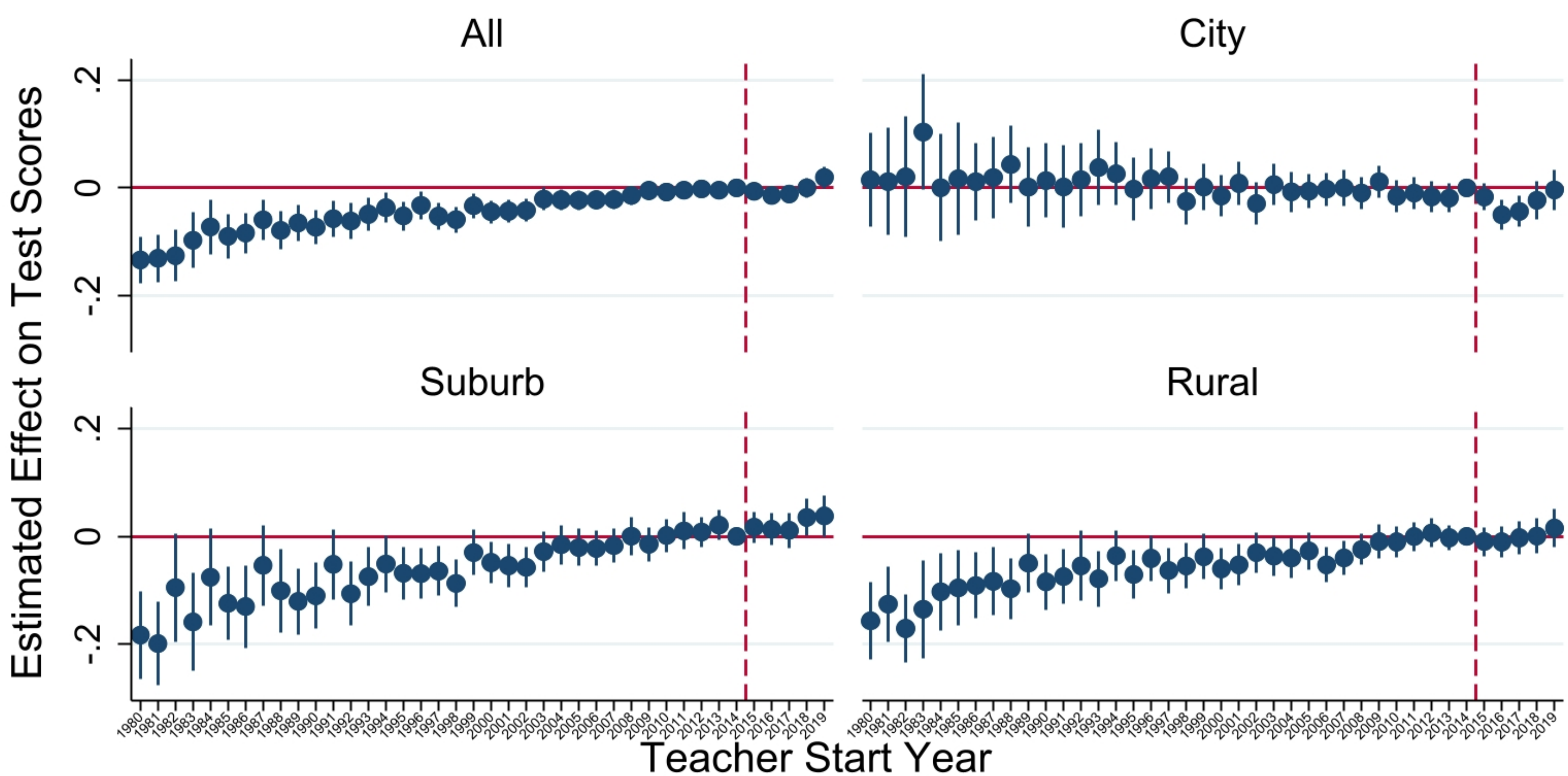
Each point shows the estimated coefficient from independent regressions on a sample restricting teacher experience level. Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include both the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed.

Figure 5: Difference in Student Test Scores By Quarter First Licensed



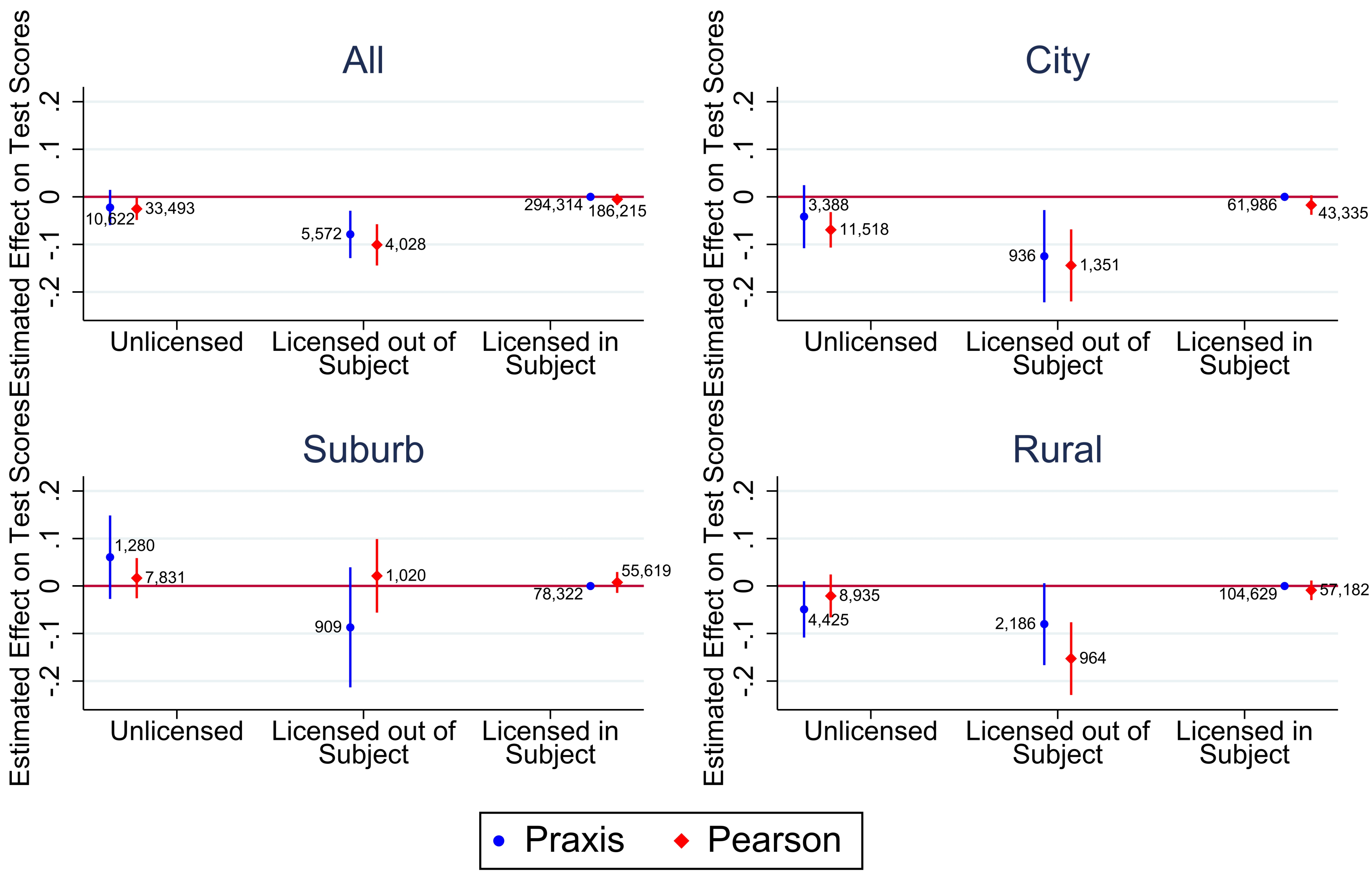
Each point shows the estimated coefficient from independent regressions on a sample restricting teacher experience level. Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include both the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. Unlicensed teachers are set to be first licensed in the third quarter of the year they begin teaching. The first dashed red line is when teachers were first able to take Pearson tests. The second red dot-dash line shows when teachers were forced to take Pearson tests rather than Praxis tests.

Figure 6: Difference in Student Test Scores By Teacher's Imputed First Year



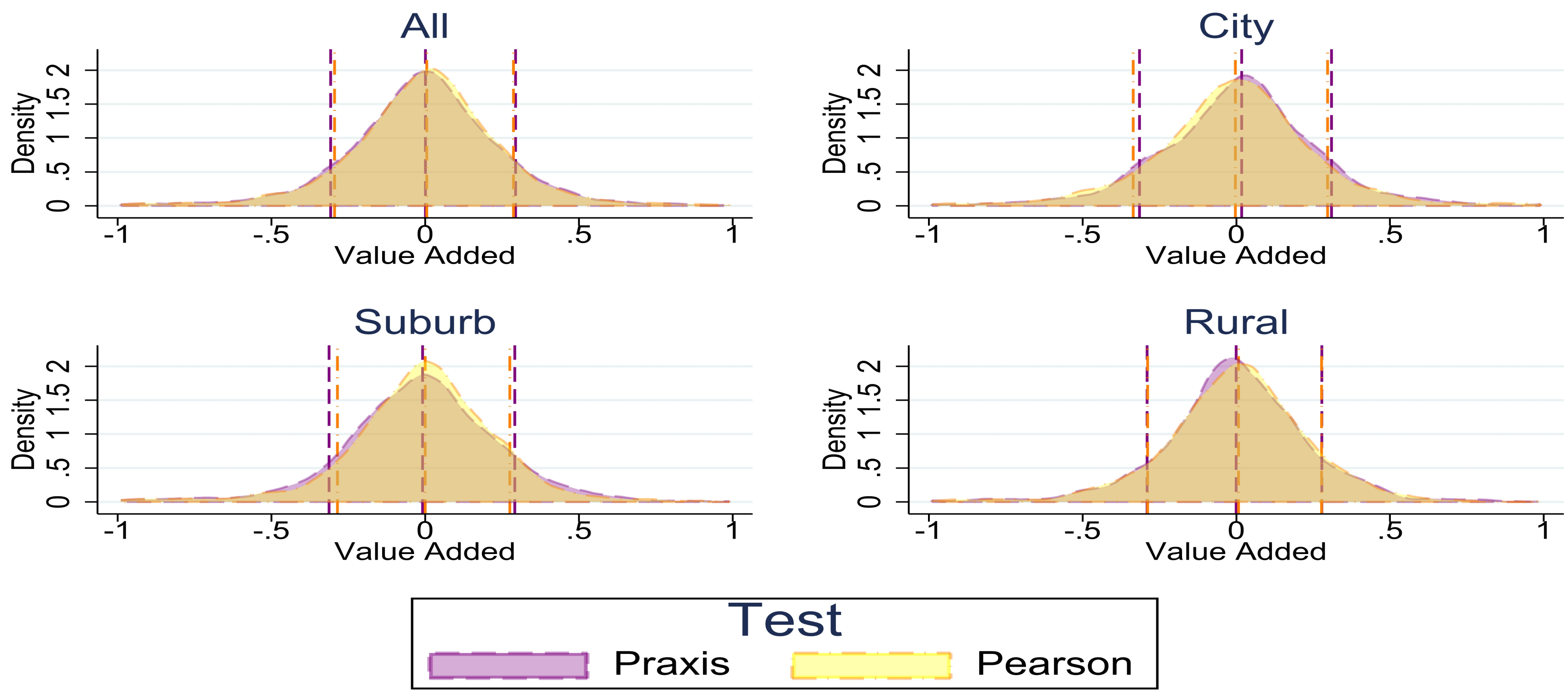
Each point shows the estimated coefficient from independent regressions on a sample restricting teacher experience level. Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include both the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. Unlicensed teachers are set to be first licensed in the third quarter of the year they begin teaching. The dashed red line is when teachers were first able to take Pearson tests. A teacher's first year is imputed by using the minimum difference between a teacher's years of experience and the current school year.

Figure 7: The Estimated Change in Teacher Value-Added by License Type



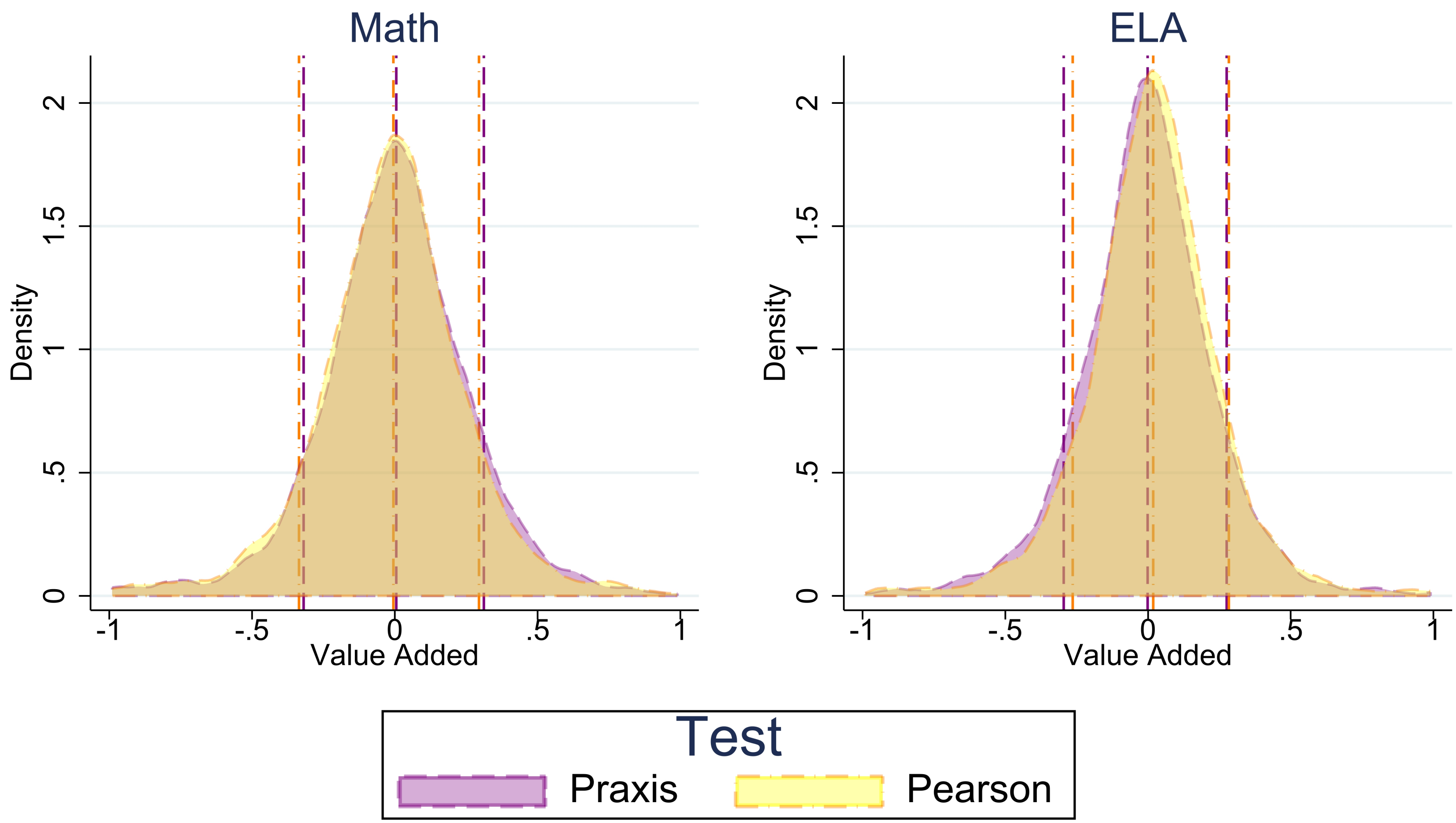
The numbers next to each point estimate signify the relevant sample sizes for each point estimate. Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include both the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. For point estimates, see Table A.1.

Figure 8: VAM Distribution Changes by Residential Setting



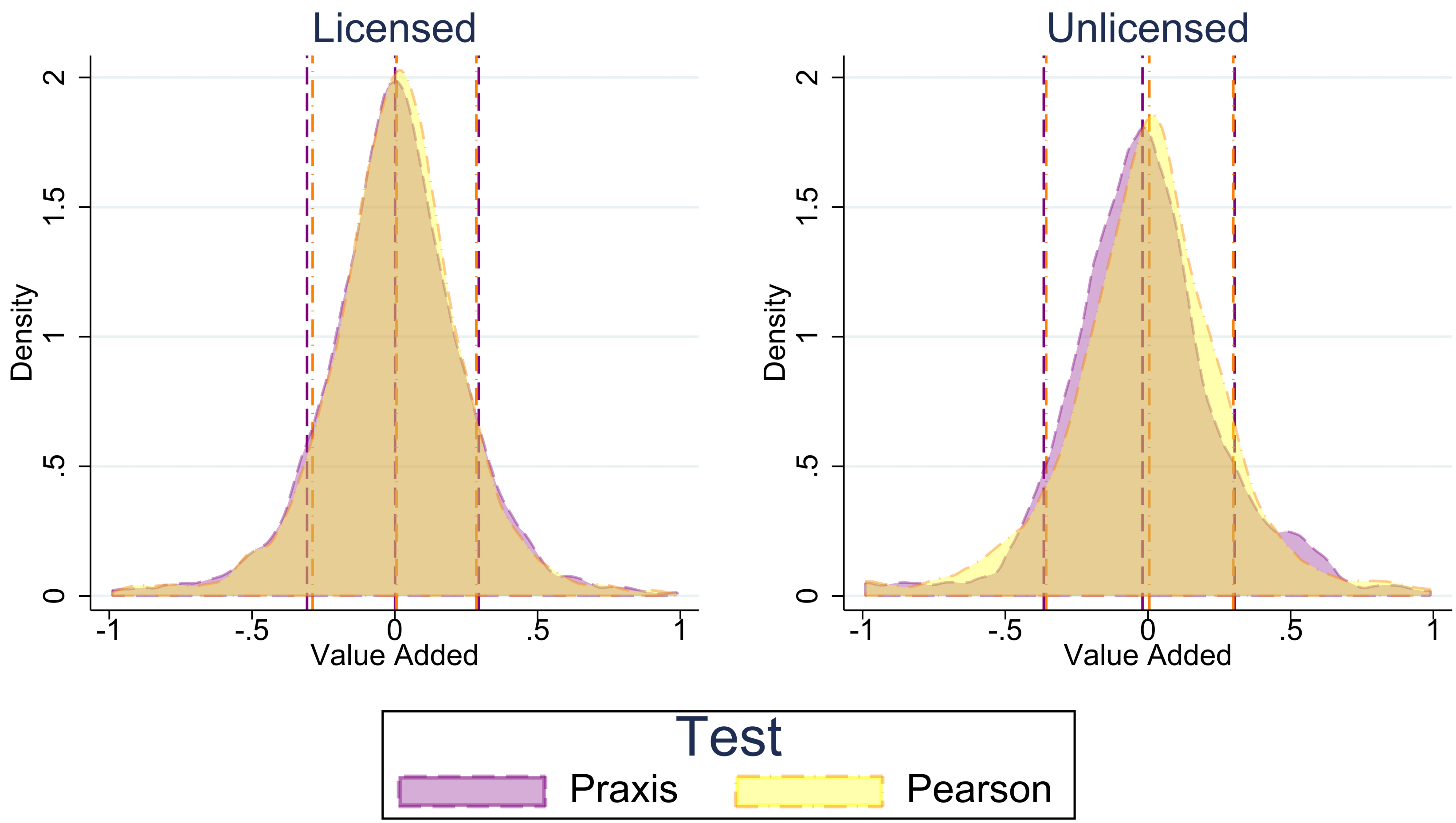
Each area represents the density plot of year-specific Value-Added for incoming teachers under Pearson and Praxis regimes, respectively centered for subject and experience level. The leftmost lines represent the 10th percentile of the distribution within each population of teachers. The middle lines demarcate the median and the rightmost demarcate the 90th percentile.

Figure 9: VAM Distribution Changes by Subject



Each area represents the density plot of year-specific Value-Added for incoming teachers under Pearson and Praxis regimes, respectively centered for subject and experience level. The leftmost lines represent the 10th percentile of the distribution within each population of teachers. The middle lines demarcate the median and the rightmost demarcate the 90th percentile.

Figure 10: VAM Distribution Changes by Licensure Status



Each area represents the density plot of year-specific Value-Added for incoming teachers under Pearson and Praxis regimes, respectively centered for subject and experience level. The leftmost lines represent the 10th percentile of the distribution within each population of teachers. The middle lines demarcate the median and the rightmost demarcate the 90th percentile.

Table 1: Summary Statistics for Test Score Analysis

	Mean	SD
Teacher Starts/Licensed Under Pearson	0.419	0.493
Student Test Score (Standardized)	-0.055	0.967
Student Ever Has IEP	0.202	0.402
Student Has Free or Reduced Price Lunch	0.490	0.500
Student is Black	0.098	0.298
Student is Hispanic	0.120	0.325
Class Subject is Math	0.483	0.500
Class Subject is English	0.517	0.500
Teacher is Unlicensed	0.083	0.275
Teacher is Licensed For Something Else	0.018	0.133
Teacher is Licensed for Class	0.899	0.301
N	540,763	

Observations are at a student-teacher-year level. These records are for students in grades 3-8, in math or English classes in the spring semesters of 2012-2019. We only include the most common class taken by a student in a subject in a given year. Test scores are standardized within grade, subject, and year before excluding observations, which is why the mean is not 0.

Table 2: The Effect of Test Regime on Student Test Scores With Responsible Teacher Randomly Selected

Regression of Test Regime on Student Test Scores			
Sample	**All**	**Schools in a City**	**Schools in a Suburb**
Estimated Effect With Standard Assignment	-0.007	-0.026*	0.009
Range With Random Assignment (Min,Max)	(-0.002, 0.002)	(-0.025,-0.017)	(0.003,0.010)
Iterations	1,000	1,000	1,000

Each regression uses the same controls as in Table 4. For each iteration under random assignment, we randomly select the teacher and the class that is responsible for a student's test score among the teachers and classes a student has within a subject and school year. *** = $p \leq .001$, ** = $p \leq .01$, * = $p \leq .05$

Table 3: Changes in New Teacher Demographics

License Type	**All TPS Teachers**		**Licensed Teachers**		**Unlicensed Teachers**	
	Praxis	**Pearson**	**Praxis**	**Pearson**	**Praxis**	**Pearson**
Highest Degree is a BA	0.926	0.933	0.938	0.952*	.863	0.875
Highest Degree is a MA	0.054	0.047	0.048	0.036*	0.115	0.079
White	0.938	0.899***	0.949	0.946	0.840	0.765
Black	0.033	0.061***	0.023	0.019	0.130	0.177
Hispanic	0.016	0.021	0.018	0.020	0.008	0.028
Male	0.176	0.189	0.166	0.164	0.260	0.263
Age	29.338	28.040***	29.346	26.996***	30.092	31.831
Licensed	0.921	0.769***				
Licensed For Math	0.113	0.069***				
Licensed For ELA	0.225	0.115***				
Licensed for SWD	0.190	0.191				
Licensed for ELLs	0.018	0.013**				
Holds Transition to Teaching License	0.025	0.038**				
Licensed for Elementary School	0.804	0.768**				
Licensed for Middle School	0.878	0.762***				
N	2,721	4,014	1994	2537	131	617

This sample only includes teachers who started teaching between 2012 and 2019. Observations are at a teacher level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2013 and are unlicensed. A Transition to Teaching license is Indiana's primary alternative pathway to teaching. *** = $p \leq .001$, ** = $p \leq .01$, * = $p \leq .05$

Table 4: The Estimated Change in Incoming Teacher Quality

Regression of Test Regime on Student Test Scores						
Sample	**All**	**Math Classes**	**ELA Classes**	**White Students**	**Students of Color**	**Free or Reduced Price Lunch**
Licensed Under Pearson	-0.007	-0.020*	0.004	-0.003	-0.019*	-0.012*
	(0.006)	(0.009)	(0.006)	(0.006)	(0.008)	(0.006)
N	257,640	276,441	380,659	153,397	264,869	216,147
Sample	**Elementary School**	**Middle School**	**Schools in a City**	**Schools in a Suburb**	**Schools in a Town**	**Schools in a Rural Area**
Licensed Under Pearson	-0.000	-0.013	-0.026*	0.009	0.007	-0.009
	(0.008)	(0.008)	(0.010)	(0.011)	(0.014)	(0.010)
N	534,244	318,075	122,514	144,981	74,683	178,321

Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include the prior year's math and ELA test scores interacted with the current year observation subject. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language learner, or special education designation change relative to the previous year. We do not control for a year fixed effect because test scores are already standardized at the grade-year-subject level. Observations are at a teacher-student-school-year level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. *** = $p \leq .001$, ** = $p \leq .01$, * = $p \leq .05$

Table 5: Quantile Regressions on Teacher Value Added

Sample	All	Licensed	Unlicensed	All	All	All					
Residential Setting	All					Urban		Suburban		Rural	
Subject	Both			Math	ELA	Math	ELA	Math	ELA	Math	ELA
10th Percentile	0.018	0.027*	0.027	-0.010	0.042**	-0.033	0.021	0.026	0.053*	-0.027	0.040
(SE)	(0.012)	(0.013)	(0.074)	(0.018)	(0.014)	(0.028)	(0.026)	(0.033)	(0.022)	(0.026)	(0.021)
Value Under Praxis	-0.308	-0.307	-0.365	-0.319	-0.294	-0.312	-0.329	-0.354	-0.288	-0.311	-0.267
50th Percentile	0.003	0.003	0.013	-0.013	0.018**	-0.056***	0.000	-0.012	0.020	-0.005	0.032**
(SE)	(0.005)	(0.006)	(0.023)	(0.008)	(0.006)	(0.015)	(0.012)	(0.013)	(0.011)	(0.011)	(0.011)
Value Under Praxis	0.001	0.001	-0.019	0.006	-0.002	0.021	0.012	0.000	-0.015	0.001	-0.004
90th Percentile	-0.011	-0.012	-0.022	-0.039**	0.005	-0.038	0.006	-0.063**	-0.007	-0.012	0.013
(SE)	(0.010)	(0.010)	(0.073)	(0.014)	(0.011)	(0.023)	(0.022)	(0.021)	(0.017)	(0.025)	(0.017)
Value Under Praxis	0.294	0.294	0.304	0.312	0.272	0.315	0.300	0.323	0.266	0.296	0.260
N	13,997	12,858	1,139	6,807	7,190	1,635	1,762	1,821	2,001	2,271	2,327

The above regressions were run on a sample of year-specific value-added measures for incoming teachers. Each value-added measure is centered for each subject and experience level. Each coefficient is the estimated difference between teachers entering under the Pearson regime rather than entering under the Praxis II regime at the reported percentile. Each regression controls for the experience level of the teacher interacted with the subject matter of the test. The "Value Under Praxis" is the value-added value at each percentile among the population that entered under the Praxis II regime.

Table 6: Estimated Changes in Teacher Quality by License Type

Regression of Test Regime on Student Test Scores						
Sample	**All**	**Math Classes**	**ELA Classes**	**White Students**	**Students of Color**	**Free or Reduced Price Lunch**
Licensed Under Pearson * Licensed in Subject	-0.005*	-0.022*	0.006	-0.001	-0.017*	-0.012
	(0.006)	(0.009)	(0.006)	(0.006)	(0.008)	(0.006)
Licensed Under Pearson * Licensed out of Subject	-0.022	0.089	-0.025	-0.033	0.012	-0.012
	(0.033)	(0.056)	(0.032)	(0.038)	(0.036)	(0.033)
Licensed Under Pearson * Unlicensed	-0.003	0.054	-0.002	0.005	-0.020	0.014
	(0.022)	(0.046)	(0.022)	(0.024)	(0.027)	(0.022
Licensed out of Subject	-0.079**	-0.203***	-0.053*	-0.074*	-0.097***	-0.101***
	(0.025)	(0.043)	(0.023)	(0.029)	(0.026)	(0.025)
Unlicensed	-0.022	-0.089*	-0.012	-0.028	-0.013	-0.046*
	(0.019)	(0.042)	(0.019)	(0.021)	(0.024)	(0.020)
N	534,244	257,640	276,441	380,659	153,397	264,869
Sample	**Elementary School**	**Middle School**	**Schools in a City**	**Schools in a Suburb**	**Schools in a Town**	**Schools in a Rural Area**
Licensed Under Pearson * Licensed in Subject	-0.003	-0.009	-0.017	0.008	0.009	-0.009
	(0.009)	(0.008)	(0.010)	(0.011)	(0.015)	(0.010)
Licensed Under Pearson * Licensed out of Subject	0.026	-0.033	-0.019	0.108	-0.135*	-0.073
	(0.061)	(0.037)	(0.061)	(0.074)	(0.063)	(0.057)
Licensed Under Pearson * Unlicensed	0.094**	-0.032	-0.028	-0.044	0.010	0.028
	(0.034)	(0.025)	(0.037)	(0.049)	(0.056)	(0.037)
Licensed out of Subject	-0.095*	-0.083**	-0.125*	-0.087	-0.005	-0.080
	(0.046)	(0.029)	(0.050)	(0.064)	(0.039)	(0.044)
Unlicensed	-0.086**	-0.009	-0.042	0.061	-0.005	-0.049
	(0.027)	(0.022)	(0.034)	(0.034)	(0.044)	(0.030)
N	216,147	318,075	122,514	144,981	74,683	178,321

Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level.We do not control for a year fixed effect because test scores are already standardized at the grade-year-subject level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. *** = $p \leq .001$, ** = $p \leq .01$, * = $p \leq .05$

References

Anderson, D. M., Brown, R., Charles, K. K., and Rees, D. I. (2020). Occupational Licensing and Maternal Health: Evidence From Early Midwifery Laws. *Journal of Political Economy*, 128(11):4337–4383.

Angrist, J. D. and Guryan, J. (2004). Teacher Testing, Teacher Education, and Teacher Characteristics. *American Economic Review*, 94(2):241–246.

Bacher-Hicks, A., Chi, O. L., Tichnor-Wagner, A., and Baloch, S. (2023). Who Becomes a Teacher when Entry Requirements Are Reduced? An Analysis of Emergency Licenses in Massachusetts. *Annenberg Institute at Brown University*.

Backes, B., Cowan, J., Goldhaber, D., and Theobald, R. (2024). Four Years of Pandemic-Era Emergency Licenses: Retention and Effectiveness of Emergency-Licensed Massachusetts Teachers Over Time. *National Center for Analysis of Longitudinal Data in Education Research*.

Backes, B. and Hansen, M. (2018). The Impact of Teach For America on Non-Test Academic Outcomes. *Education Finance and Policy*, 13(2):168–193.

Bhattacharya, V., Illanes, G., and Padi, M. (2019). Fiduciary Duty and the Market for Financial Advice. Working Paper 25861, National Bureau of Economic Research.

Carroll, S. L. and Gaston, R. J. (1981). Occupational Restrictions and the Quality of Service Received: Some Evidence. *Southern Economic Journal*, 47(4):959–976.

Chung, B. W. and Zou, J. (2022). Teacher Licensing, Teacher Supply, and Student

Achievement: Nationwide Implementation of edTPA. *Annenberg Institute at Brown University*, (440).

Clark, M. A. and Isenberg, E. (2020). Do Teach For America Corps Members Still Improve Student Achievement? Evidence from a Randomized Controlled Trial of Teach For America's Scale-Up Effort. *Education Finance and Policy*, 15(4):736–760.

Clotfelter, C. T., Ladd, H., and Vigdor, J. L. (2007a). How and Why Do Teacher Credentials Matter for Student Achievement? Working paper.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2006). Teacher-Student Matching and The Assessment of Teacher Effectiveness. *Journal of Human Resources*, 41(4):778–820.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2007b). Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects. *Economics of Education Review*, 26(6):673–682.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2010). Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects. *Journal of Human Resources*, 45(3):655–681.

Cowan, J., Goldhaber, D., Jin, Z., and Theobald, R. (2020). Teacher Licensure Tests: Barrier or Predictive Tool? Working Paper No. 245-1020. *National Center for Analysis of Longitudinal Data in Education Research (CALDER).*

DeArmond, M., Goldhabe, D., and Payne, S. (2023). COVID's Under-the-Radar Experiment with Teacher Licensure. CALDER Research Brief. *National Center for Analysis of Longitudinal Data in Education Research (CALDER).*

Farronato, C., Fradkin, A., Larsen, B., and Brynjolfsson, E. (2020). Consumer Protection in an Online World: An Analysis of Occupational Licensing. Working Paper 26601, National Bureau of Economic Research.

Goldhaber, D. (2007). Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness? *Journal of Human Resources*, 42(4):765–794.

Hall, J., Hicks, J., Kleiner, M. M., and Solomon, R. (2019). Occupational Licensing of Uber Drivers. *Stanford Institute for Theoretical Economics*.

Hanushek, E. A. and Pace, R. R. (1995). Who Chooses to Teach (and Why)? *Economics of Education Review*, 14(2):101–117.

Henry, G. T., Bastian, K. C., Fortner, C. K., Kershaw, D. C., Purtell, K. M., Thompson, C. L., and Zulli, R. A. (2014). Teacher Preparation Policies and Their Effects on Student Achievement. *Education Finance and Policy*, 9(3):264–303.

Indiana Department of Education (2023). Assingnment Code.

Kleiner, M. M. and Kudrle, R. T. (2000). Does Regulation Affect Economic Outcomes? The Case of Dentistry. *The Journal of Law and Economics*, 43(2):547–582.

Kleiner, M. M. and Soltas, E. J. (2019). A Welfare Analysis of Occupational Licensing in US States. Working Paper 26383, National Bureau of Economic Research.

Kugler, A. D. and Sauer, R. M. (2005). Doctors Without Borders? Relicensing Requirements and Negative Selection in the Market for Physicians. *Journal of Labor Economics*, 23(3):437–465.

Larsen, B., Ju, Z., Kapor, A., and Yu, C. (2020). The Effect of Occupational Licensing Stringency on The Teacher Quality Distribution. Technical report, National Bureau of Economic Research.

Master, B. K., Schweig, J., Berglund, T., and Hu, L. (2023). Teaching and Leading: An Evaluation of the Impacts and Professional Trajectories of Teach for America Participants in Indianapolis. Research Report. RR-A2127-1. *RAND Corporation*.

Olivia L. Chi, Andrew Bacher-Hicks, A. T.-W. S. B. (2024). Teacher Licensure and Workforce Quality: Insights from Covid-Era Emergency Licenses in Massachusetts. Technical Report 936, Annenberg Institute at Brown University.

Penner, E. K. (2021). Teach For America and Teacher Quality: Increasing Achievement Over Time. *Educational Policy*, 35(7):1047–1084.

Ramseyer, J. M. and Rasmusen, E. B. (2015). Lowering the Bar to Raise the Bar: Licensing Difficulty and Attorney Quality in Japan. *The Journal of Japanese Studies*, 41(1):113–142.

Segall, B. (2017). Crisis in the Classroom: New Indiana Teachers Repeatedly Failing State Exams — wthr.com.

Shuls, J. V. (2018). Raising The Bar on Teacher Quality: Assessing The Impact of Increasing Licensure Exam Cut-Scores. *Educational Policy*, 32(7):969–992.

Shuls, J. V. and Trivitt, J. R. (2015). Teacher Effectiveness: An Analysis of Licensure Screens. *Educational Policy*, 29(4):645–675.

Slay, L. E., Riley, J., and Miller, K. (2020). Facilitating a Path to New Teacher Certification

Amid the COVID-19 Pandemic: Unpacking States' "Unchanged-New Flex" guidelines. In *Frontiers in Education*, volume 5, page 583896. Frontiers Media SA.

Strauss, R. P. and Sawyer, E. A. (1986). Some New Evidence on Teacher and Student Competencies. *Economics of Education Review*, 5(1):41–48.

Xu, Z., Hannaway, J., and Taylor, C. (2011). Making a Difference? The Effects of Teach for America in High School. *Journal of Policy Analysis and Management*, 30(3):447–469.

A Heterogeneity By License Type

Table A.1: Estimated Changes in Teacher Quality by License Type

Regression of Test Regime on Student Test Scores						
Sample	**All**	**Math Classes**	**ELA Classes**	**White Students**	**Students of Color**	**Free or Reduced Price Lunch**
Licensed Under Pearson * Licensed in Subject	-0.005*	-0.022*	0.006	-0.001	-0.017*	-0.012
	(0.006)	(0.009)	(0.006)	(0.006)	(0.008)	(0.006)
Licensed Under Pearson * Licensed out of Subject	-0.022	0.089	-0.025	-0.033	0.012	-0.012
	(0.033)	(0.056)	(0.032)	(0.038)	(0.036)	(0.033)
Licensed Under Pearson * Unlicensed	-0.003	0.054	-0.002	0.005	-0.020	0.014
	(0.022)	(0.046)	(0.022)	(0.024)	(0.027)	(0.022
Licensed out of Subject	-0.079**	-0.203***	-0.053*	-0.074*	-0.097***	-0.101***
	(0.025)	(0.043)	(0.023)	(0.029)	(0.026)	(0.025)
Unlicensed	-0.022	-0.089*	-0.012	-0.028	-0.013	-0.046*
	(0.019)	(0.042)	(0.019)	(0.021)	(0.024)	(0.020)
N	534,244	257,640	276,441	380,659	153,397	264,869
Sample	**Elementary School**	**Middle School**	**Schools in a City**	**Schools in a Suburb**	**Schools in a Town**	**Schools in a Rural Area**
Licensed Under Pearson * Licensed in Subject	-0.003	-0.009	-0.017	0.008	0.009	-0.009
	(0.009)	(0.008)	(0.010)	(0.011)	(0.015)	(0.010)
Licensed Under Pearson * Licensed out of Subject	0.026	-0.033	-0.019	0.108	-0.135*	-0.073
	(0.061)	(0.037)	(0.061)	(0.074)	(0.063)	(0.057)
Licensed Under Pearson * Unlicensed	0.094**	-0.032	-0.028	-0.044	0.010	0.028
	(0.034)	(0.025)	(0.037)	(0.049)	(0.056)	(0.037)
Licensed out of Subject	-0.095*	-0.083**	-0.125*	-0.087	-0.005	-0.080
	(0.046)	(0.029)	(0.050)	(0.064)	(0.039)	(0.044)
Unlicensed	-0.086**	-0.009	-0.042	0.061	-0.005	-0.049
	(0.027)	(0.022)	(0.034)	(0.034)	(0.044)	(0.030)
N	216,147	318,075	122,514	144,981	74,683	178,321

Each regression includes class size, grade, subject, and experience level-fixed effects. These regressions use school years 2013-2019 and grades 4-8. Controls include the math and ELA test scores from the prior year interacted with the subject of the current year observation. Controls also include the student's demographics (gender, ethnicity, English language learner status, and special education status). Schools in Indiana appear to have differing patterns of free and reduced-price lunch (FRPL) assignments, so we include school-by-FRPL fixed effects. We also control for if a student has their FRPL, English language, or special education designation change relative to the previous year. Observations are at a teacher-student-school-year level.We do not control for a year fixed effect because test scores are already standardized at the grade-year-subject level. A teacher is "Licensed Under Pearson" if their license starts in 2014 or later or if they are teaching after 2014 and are unlicensed. *** = $p \leq .001$, ** = $p \leq .01$, * = $p \leq .05$

B Value-Added Calulation

We calculate our value-added measure by regressing on student test scores using a fixed effect for a teacher in a subject in a year (δ_{tsy}). We control for student assignment by controlling for a cubic polynomial of student test scores in the prior year, the demographics of the test taker, the demographics of the full student body taught by a teacher, as well as subject and school-level fixed effects (X_{tcpsy}).

$$Y_{icgsjt} = \alpha + \chi X_{icsjt} + \gamma f(Y_{icsjt-1}) + \delta_{tcj} + \lambda_c + \psi_g + \varepsilon_{icgsjt} \tag{3}$$

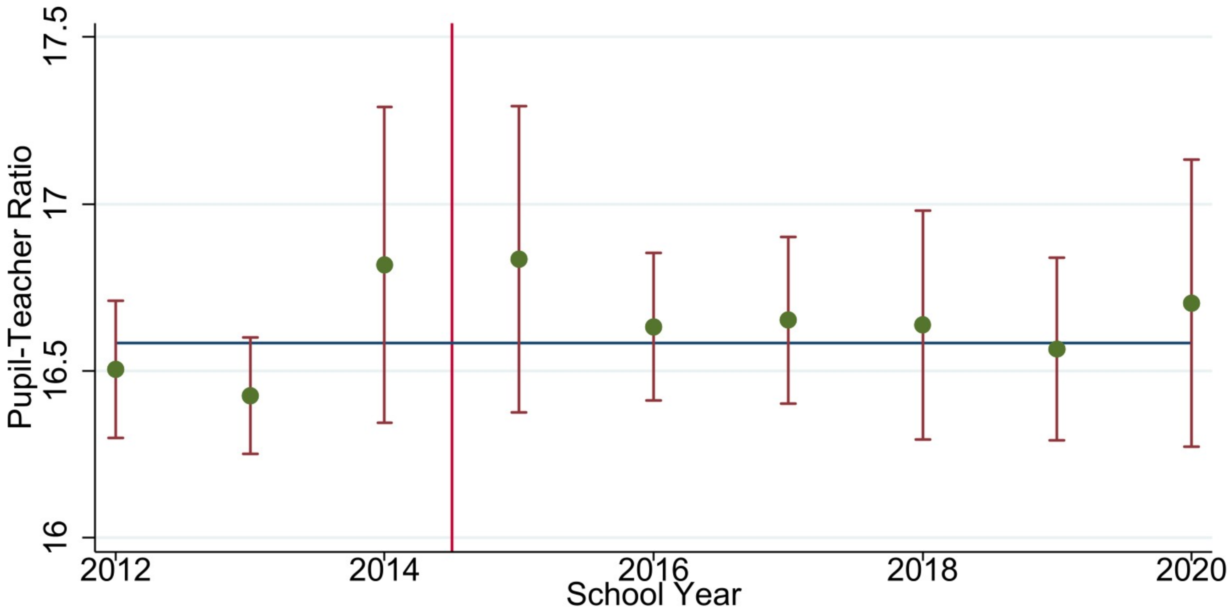
We then regress the test regime a teacher entered under on our calculated value-added measure. We test for changes at the 10th, 50th and 90th percentiles.

$$\delta_{tcj} = \phi PEARSON_j + \sum_{i=0}^{4} \theta_i English_c * 1[EXP_{jt} = i] + \mu_{tcj} \tag{4}$$

$PEARSON_j$ is a dummy variable for if a teacher was licensed after the switch to the Pearson Core tests. $\sum_{i=0}^{4} \theta_i English_c * 1[EXP_{tj} = i]$ are controls for the subject by experience-level of the teacher.

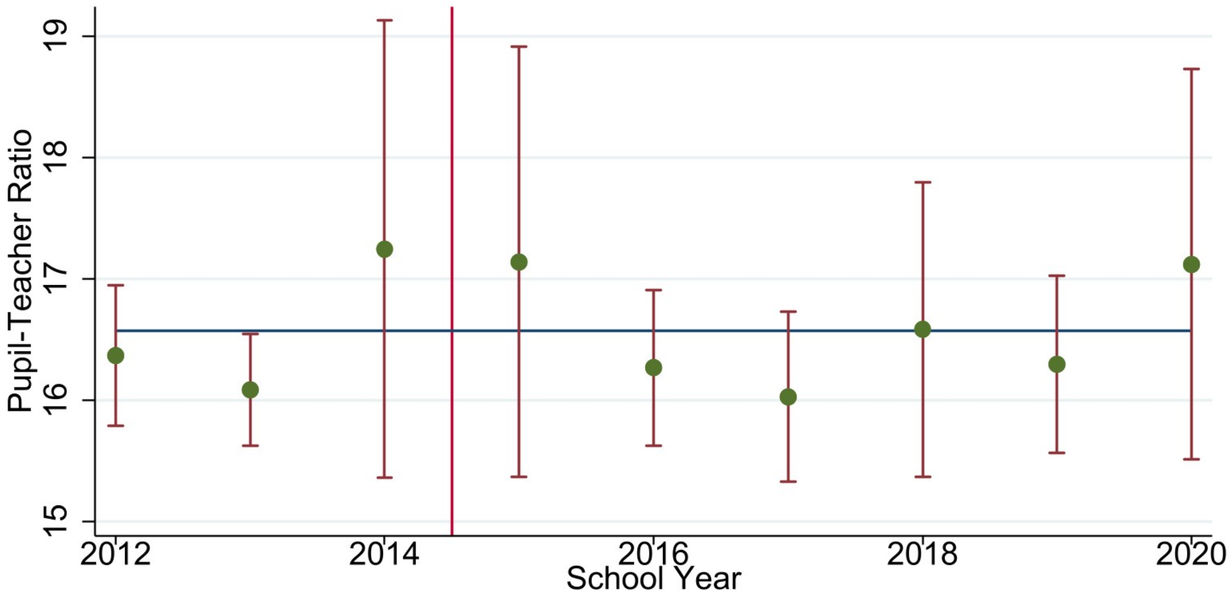
C Pupil-Teacher Ratio by School Types

Figure C.1: Pupil-Teacher Ratio at Traditional Public Schools by Year



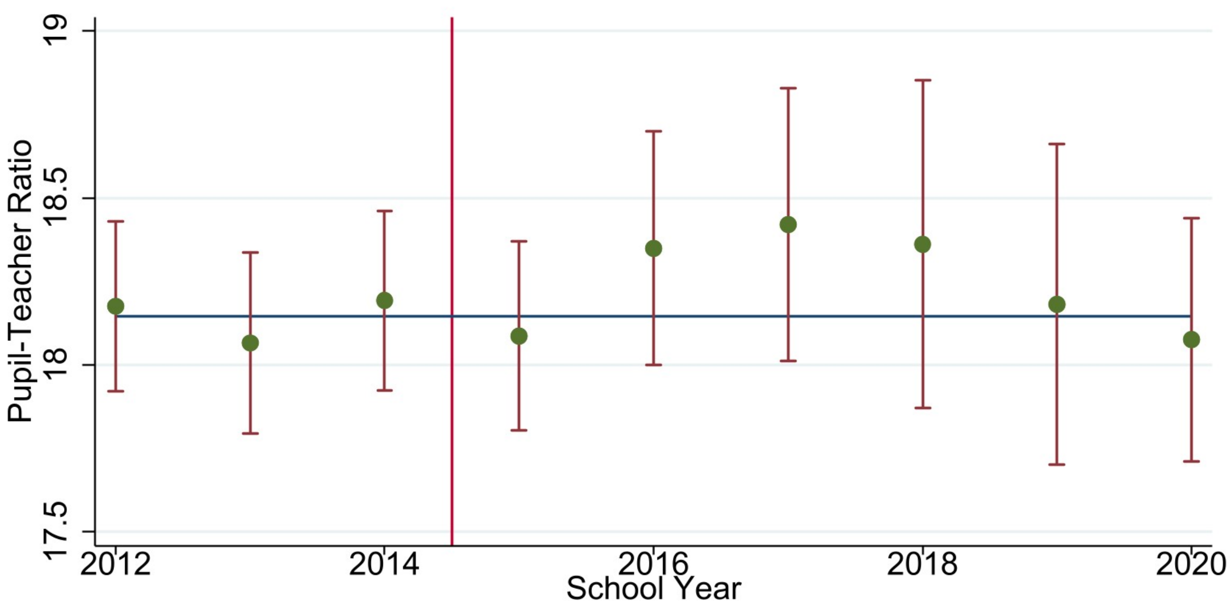
Pupil-Teacher Ratio is calculated as a fraction of students enrolled over the number of teachers assigned to a class in a school and in a given school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.

Figure C.2: Pupil-Teacher Ratio at City Schools by Year



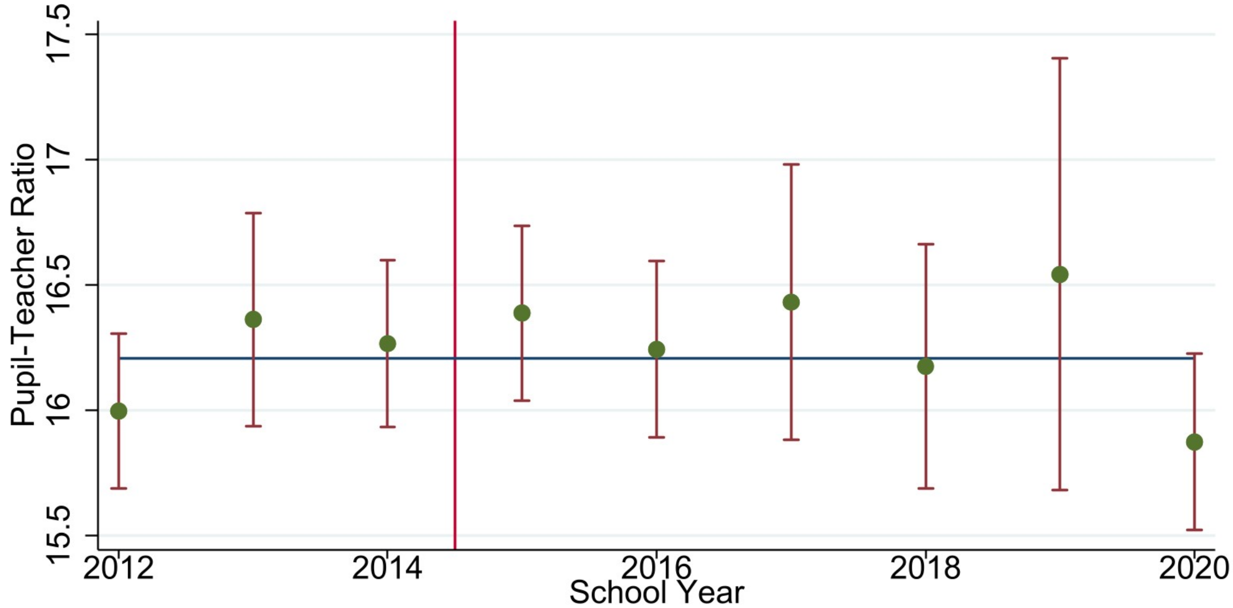
Pupil-Teacher Ratio is calculated as a fraction of students enrolled over the number of teachers assigned to a class in a school and in a given school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.

Figure C.3: Pupil-Teacher Ratio at Suburban Schools by Year



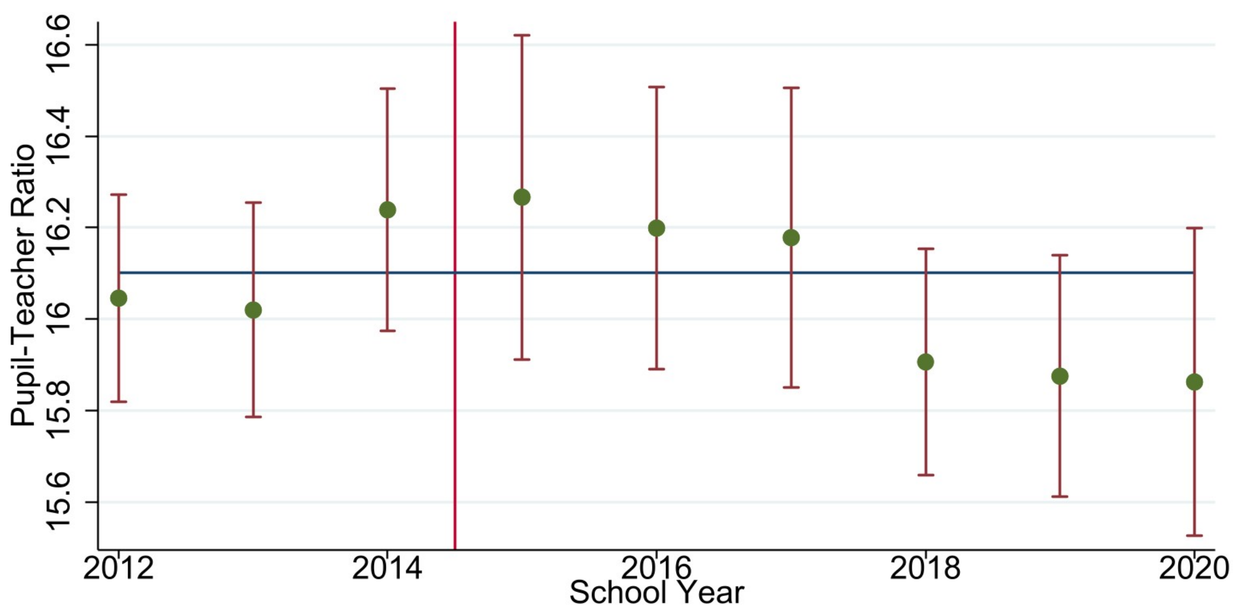
Pupil-Teacher Ratio is calculated as a fraction of students enrolled over the number of teachers assigned to a class in a school and in a given school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.

Figure C.4: Pupil-Teacher Ratio at Town Schools by Year



Pupil-Teacher Ratio is calculated as a fraction of students enrolled over the number of teachers assigned to a class in a school and in a given school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.

Figure C.5: Pupil-Teacher Ratio at Rural Schools by Year



Pupil-Teacher Ratio is calculated as a fraction of students enrolled over the number of teachers assigned to a class in a school and in a given school year. The bars represent the 95 % confidence interval for each point. The line is the mean for teachers in the initial Praxis era. Effectively, the error bars are performing a t-test for each year relative to the average of the first two years.