Teacher Testing Standards and the New Teacher Pipeline*

Marc T. Law† Mindy S. Marks‡ Tomer Stern§

7 September 2025

Abstract

We examine the impact of changes in testing standards on the new teacher pipeline by exploiting a quasi-experimental change in the minimum passing score required for admission into teacher training programs. Using institution-level data, we find that higher testing standards sharply reduce teacher preparation program enrollments and graduations. Increases in testing standards are also associated with declines in the number of new teacher licenses awarded. According to our estimates, the average increase in standards led to a 9 percent decrease in graduations from teacher training programs, accounting for more than one-third of the overall decline in teacher training graduates.

Keywords: testing, teacher supply, entry standards, occupational licensing, teacher shortage
JEL codes: J20, J44, J45, J48, I23

*For helpful comments on earlier drafts of this paper we would like to thank Emily Beam, Bobby (Wing Yin) Chung, seminar participants at Northeastern University and the University of Southern Denmark, and participants at the Fall 2023 APPAM research conference. The authors have no relevant conflicts of interest to disclose.

†Department of Economics, University of Vermont, Burlington, VT, USA (email: marc.law@uvm.edu)

‡Department of Economics, Northeastern University, Boston, MA, USA (email: m.marks@northeastern.edu)

§Burning Glass Institute, New York, NY, USA (email: tomerstern13@gmail.com)

1 Introduction

Testing is ubiquitous in settings with asymmetric information about quality. Before gaining the right to practice, prospective doctors and lawyers must pass licensing exams. Firms sometimes ask job applicants to take skills tests prior to being hired. In the context of US undergraduate programs, admissions committees use Scholastic Aptitude Test (SAT) or American College Test (ACT) scores to screen out applicants. The primary objective of these entry standards is to ensure a minimum level of quality among entrants. However, there is an inherent tradeoff: while rigorous standards help maintain quality, they may also reduce the size of the entrant pool. Since testing is imperfect, raising entry standards may inadvertently exclude some individuals who would have succeeded, thereby shrinking the pool of qualified candidates.[1] Despite its widespread use, empirical evidence on how testing impacts entrants remains limited. This gap is partly because testing standards are often adjusted endogenously in response to market conditions, making it challenging to causally identify their impact.

In this paper, we investigate how testing standards affect the progression of individuals through the professional pipeline. Our setting is the US teaching profession, one of the nation's largest occupations, and one in which state governments require potential teachers to pass a variety of tests at various stages along the new teacher pipeline, including entry into university teacher preparation programs. We take advantage of a unique quasi-experiment that provides exogenous variation in difficulty of entry testing standards for potential teachers, namely the replacement of the Educational Testing Service's (ETS) Pre-Professional Skills Test (PPST) with the Praxis Core.[2] The PPST and the Praxis Core are a series of tests that prospective teachers in many states must pass prior to admission into university teacher preparation programs. The two tests cover the same subjects (basic reading, writing, and mathematics) and are similar in terms of topics, length, and structure. Under the PPST,

[1] Moreover, in a world of imperfect information about performance on an entry test, higher entry standards may discourage potentially successful candidates from applying. See Guasch and Weiss (1981).

[2] The PPST and Praxis Core are sometimes referred to as the Praxis I exams.

the minimum passing scores set by each state varied widely. Because the Praxis Core was scaled differently than the PPST, new passing scores had to be set with limited information about the difficulty of the new test. To facilitate this, the ETS organized a multi-state standard-setting panel to determine recommended passing scores for the new exam, which were adopted by every state but one. Replacement of the PPST with the Praxis Core raised entry standards on average, but the extent of change differed widely across states. We argue that the extent to which the replacement of the PPST with the Praxis Core changed the difficulty of the entry standards in each state should be uncorrelated with factors that affected the demand and supply of teachers, given that almost every state adopted the same passing score in spite of vastly different standards under the PPST and different conditions in each state.

To investigate the impact of testing on supply, we analyze institution-level data on enrollments and graduations from university teacher training programs in 21 states plus the District of Columbia that replaced the PPST with the Praxis Core. Both outcomes are examined because an increase in entry standards could reduce enrollments without affecting graduations, leaving the supply of new teachers unchanged. Holding graduation standards constant, higher testing standards do not reduce the supply of new teachers if the decline in enrollments only eliminates students who would have been unable to complete the program. However, if the higher standards exclude candidates who would have successfully graduated, both enrollments and graduations will decrease, ultimately reducing the supply of new teachers.[3]

We find that increases in testing standards induce an economically significant reduction in the number of people entering teacher preparation programs. Two-way fixed effect estimates indicate that a one-standard deviation increase in testing difficulty decreases enrollment in

[3]One concern with increased entry standards is that organizations will endogenously increase exit standards in response to higher quality candidates. If this occurs, we will be unable to isolate the impact of changing entry standards. An advantage of our setting is the fact that most education programs operate under state-mandated exit requirements. Specifically, in addition to entry exams, teachers must pass standardized exit exams. Importantly, these exit requirements were unchanged during our sample period.

education programs by 16 percent and graduations by 22 percent. Using an event-study approach, we estimate that the increase in standards in 2013 reduced enrollments in education majors instantaneously and reduced graduations from teacher preparation programs with a lag. The timing of these effects is consistent with the fact that prospective teachers typically take these entry exams two or three years prior to their graduation. The fact that the pre-trends are small and statistically indistinguishable from zero increases confidence that the change in testing standards was orthogonal to underlying trends in teacher program enrollment. Our baseline findings are robust to using measures of teacher training program enrollment and graduations from an alternative data set. Finally, we also show that the increase in testing standards is associated with a reduction in the number of new teacher licenses awarded by state governments and an increase in reported teacher shortage areas.

We contribute to the literature on teacher licensing requirements. Much of this literature examines the impact of exit tests (i.e., tests taken after teacher preparation program completion but prior to certification, for instance, the Praxis II tests) on student performance (e.g., Buddin and Zamarro, 2009; Clotfelter et al., 2010; Goldhaber, 2007; Goldhaber and Hansen, 2010; Goldhaber et al., 2017; Chen et al., 2023).[4] While the evidence is mixed, many of these studies find that teachers with higher exit test scores improve students' performance in math. Overall, teacher testing appears to raise teacher quality by increasing attrition among low-performing teachers.

A much smaller literature examines the impact of standards on the new teacher pipeline. Chung and Zou (2025) investigate the impact of the state-level adoption of the edTPA program—a portfolio based teacher evaluation program that some states added as a requirement for education program completion—on the supply of new teachers. They find that higher exit standards arising from the staggered adoption of edTPA reduce the number of graduates from teacher training programs. Like us, Chung and Zou (2025) explore the im-

[4]A related literature examines the impact of alternative teacher certification pathways on student achievement (Decker et al., 2012; Jacob and Lefgren, 2004; Kane et al., 2008; Sass, 2015). In most states, teachers obtaining alternative certifications must pass the required exams but are exempt from the coursework that is taken in conventional teacher preparation programs.

pact of a requirement embedded within teacher training programs. However, edTPA is a semester-long project involving lesson plans, classroom videos, and follow-up reports whereas our setting is a basic skills test. Additionally, the set of states adopting edTPA differs from those affected by our experiment. In related work, Orellana and Winters (2025) document that many students fail the Praxis exams in Connecticut. Using a regression discontinuity design they show that failing the first attempt at the Praxis Core or the Praxis II reduces the likelihood of obtaining a teaching certification. These findings lend credence to our hypothesis that testing standards are a binding constraint. Our findings generalize this insight to a broader sample of states.

Our paper is also related to Kraft et al. (2020) who examine the impact of statewide teacher accountability reforms. Although these reforms were designed to identify and remove low-performing teachers, Kraft et al. (2020) show that they also reduce the supply of new teachers. However, while accountability reforms operate in the classroom and indirectly influence entry through expectations about future working conditions, our setting—entry exams—influence who enters the pipeline in the first place.

We also connect to the literature that examines teacher preparation entry tests (i.e., the PPST), but these studies generally focus on the effect of testing standards on teacher quality. Angrist and Guryan (2008), for instance, find that requiring teachers to take the PPST raises teacher wages, but has no impact on measures of teacher input quality (e.g., teacher SAT scores). Larsen et al. (2020) incorporate the PPST exams into a broader index of licensing stringency and find that stricter licensing requirements may be effective in screening out less qualified candidates from the teaching profession (where quality is measured by the selectivity of a candidate's undergraduate institution). Shuls (2018) documents that teachers who pass the PPST did not raise student test scores relative to teachers who failed but were still permitted to teach within the state Arkansas.

Our findings also contribute to the growing literature on the consequences of occupational licensing laws on labor markets (e.g., Kleiner, 2006, 2013; Kleiner and Krueger, 2010). Much

of this literature looks at the impacts of changing licensing standards on wages (e.g., Law and Marks, 2017; Gittleman et al., 2018; Kleiner and Krueger, 2013; Thornton and Timmons, 2013). Other studies analyze the impacts of licensing laws on participation and sometimes find that tougher licensing requirements reduce participation in regulated occupations (e.g., Pagliero, 2010; Barrios, 2022; Blair and Chung, 2019; Law and Marks, 2009; Law and Kim, 2005). These papers focus on testing, training and fee regulations that apply before gaining the right to practice (but after formal training has ended). In contrast we examine requirements that are imposed before formal training begins. Accordingly, our focus is on a salient requirement that affects an earlier segment of the new practitioner pipeline. While many occupational licensing regimes require prospective practitioners to pass a test or series of tests as a prerequisite for entry, the impact of testing standards is relatively under-studied, perhaps because standards can be adjusted endogenously to demand and supply conditions (Pagliero, 2013).

Finally, our findings suggest that increases in teacher testing standards may be an important but overlooked factor behind ongoing teacher shortages. Between the 2008-09 and 2018-19 academic years, graduations from traditional teacher training programs—the primary pathway into the teaching profession—declined by 35 percent nationally, with the states we analyze also experiencing large reductions in teacher preparation program graduations (Will, 2022). This has led to much soul-searching about a crisis in the teaching profession and speculation that recent teacher shortages are due to low teacher salaries, the lack of respect afforded teachers, and new challenges in the classroom, for instance, larger class sizes, evaluation processes that depend on standardized testing in the classroom, and an increasingly politicized classroom environment (Schmitt and deCourcy, 2022). Curiously, however, the role of teacher testing is neglected in these discussions. Back-of-the-envelope calculations using our estimates suggest that the increase in testing standards arising from the replacement of the PPST with the Praxis Core accounts for more than one third of recent declines in new teacher graduations and more than half of the decline in the states where

testing standards increased the most. While we remain agnostic about the impact of these changing standards on teacher quality or student achievement, our findings indicate that rising standards play an economically significant role in reducing the supply of new teachers.

2 Background

All states require teachers to possess a license prior to teaching in public schools, with the licensing requirements set at the state-level. State standards may include requirements for entry into teacher preparation programs, course work requirements, student teaching hours, as well as exit exams on subject matter and pedagogy. Prior to becoming fully fledged teachers, prospective teachers typically enroll in a state-recognized teacher training program housed within an institution's education school, although alternative pathways are available in many states.

A universal theme in teacher licensing is that all prospective teachers must pass tests at various points along the pipeline. A subset of states requires prospective teachers to pass a series of entry tests, normally taken during their sophomore or junior year, to enter their college's teacher preparation program.[5] Some states (e.g., California) have designed their own entry tests. Others outsource this task to the ETS, which set a series of tests called the Pre-Professional Skills Test (PPST).[6] In the spring of 2013, ETS introduced the Praxis Core Academic Skills Test as a replacement for the PPST, which was fully phased-out on September 1st of 2014. Both the PPST and Praxis Core test basic knowledge of math, reading, and writing. They each consist of three component tests which can be taken all at once or on separate occasions. Prospective teachers must obtain a minimum score on each component test and/or a minimum composite score across the three tests in order to gain entry to teacher preparation programs.[7]

[5]For students entering masters' level teacher preparation programs, passing these entry exams is often required for admission.

[6]Between 2007-2010, roughly 150,000 undergraduates took each of the three the PPST exams (ETS, 2010)

[7]Many states exempt students with sufficient high scores on the ACT, SAT, or GRE from having to

According to ETS (2013), the new test was designed to align with the Common Core. Unlike the PPST, the Praxis Core is only offered as a computer-based test. While the test still assessed basic skills in the same three subject areas—reading, writing and mathematics—the questions were designed to be more focused on analyzing, assessing and evaluating various pieces of content. Additionally, the format was designed to be more interactive. There are some differences in the number of questions and the time allotted to each component. However, the overall content is remarkably similar (see Appendix Figure A1 for a comparison of the content of the tests). A comparison of the study guides provided by the ETS for both tests shows significant overlap in practice questions. For our purposes, the key difference between the two tests is that the PPST was scored from 150-190 while the Praxis Core is scored from 100-200, which required state governments to set new passing scores.

The PPST had no nationally recommended passing scores; every state set its own passing scores for each PPST component. However, when the ETS was planning for the replacement of the PPST, it convened a multi-state panel in February 2013 to establish a universal recommended passing score for the reading, writing, and mathematics components of the Praxis Core. Education agencies from 23 states, Washington D.C., and Guam sent panelists familiar with what is required of students to succeed in teacher preparation programs to recommend new passing scores. As shown in Table 1 the states that participated in the panel (Column 5) overlap heavily with our sample of states (Column 6).

3 Data and sample

Information on which entry tests are required of prospective teachers was obtained from ETS and cross-checked with state Department of Education websites. Information about the use of the PPST, the adoption of the Praxis Core, and participation in the multi-state

take Praxis entry tests. However, the threshold scores for this exemption are high. For instance, Maryland requires a minimum ACT composite score of 24 (74 percentile of all college-going students) or a total math and critical reading score of 1180 on the SAT whereas the minimum passing score for the Praxis Core is much lower at almost half a standard deviation below the national average of potential teachers.

panel is shown in Table 1. To be included in our sample, a state had to use the PPST from 2008-2012 and adopt the passing scores recommended by the multi-state panel for the Praxis Core. Four states (Iowa, Kentucky, Rhode Island, Washington) were excluded because while they adopted the Praxis Core they did not use the PPST. Indiana was excluded because it used the PPST but did not adopt the Praxis Core. Among the states that used the PPST and adopted the Praxis Core, every state but North Dakota adopted the passing scores recommended by the panel;[8] Oklahoma allowed candidates to take the Oklahoma General Education Test in place of the Praxis Core; and sources disagree about whether Tennessee required the Praxis Core.[9] Accordingly, we exclude these three states from the main sample. This leaves us with a sample of 21 states plus Washington, DC.

During our sample period, four states weakened their Praxis Core requirements. In the fall of 2016, Connecticut eliminated the requirement that students pass the Praxis Core. Also in 2016, Pennsylvania lowered the math passing score from 150 to 142; additionally, Maine changed its requirements to allow test takers to achieve a combined score of 468 with no single score lower than three points lower than the minimum passing score. Finally, in June 2017 South Carolina lowered its math and writing cutoff scores retroactively to September 2016.[10]

We measure the potential teacher supply response to the change in the entry test requirement along two margins: the number of students enrolled in education majors in post-secondary institutions, and the number of graduates from teacher preparation programs. Both data series were obtained from the Integrated Postsecondary Education Data Sys-

[8]North Dakota adopted the recommended passing score for the math and reading test but chose a lower score on the writing test. However, as shown in Appendix Table A1, the point estimates of interest are similar when we include North Dakota in the sample.

[9]The ETS reports that Tennessee requires the Praxis Core while the Tennessee Department of Education states, "Praxis Core academic tests may be required for admission to an educator preparation program but may not be required for licensure." See tn.gov (Accessed: November 27, 2023). However, our findings are robust to the inclusion of Tennessee (see Appendix Table A1).

[10]Our two-way fixed effect specifications account for these changes in passing scores and exclude Connecticut from the later years of the sample. However, as shown in column (5) of Appendix Table A1 the findings are robust to excluding this variation (which is likely endogenous). Our event study results ignore all post-period variation in testing standards.

tem (IPEDS), which contains detailed information about post-secondary institutions located within all 50 states plus DC.

In September 2019 Praxis changed the Core math test and recommended new passing scores. Accordingly, our experiment ends in 2019. Our enrollment analysis encompasses the years 2008 to 2018, since the 2018 entrants were the last available group covered by this experiment (the enrollment data is biennial; thus the 2020 entrants would have faced different passing scores and a different test). Because the PPST and Praxis Core are entry exams that are typically taken in the sophomore or junior years or prior to admission into a master's level program, the effect on graduation will occur with a lag of at least one year. Our graduation analysis therefore covers the years 2009 to 2020.

In IPEDS, fall enrollment data is available at the two-digit CIP code level for even-numbered years. We measure enrollments using CIP code 13, which is a count of the number of education students at the institution-level. This is an imprecise measure of the potential teacher pipeline. While CIP code 13 includes students in teacher preparation programs, it also includes students majoring in fields like school counseling services or instructional education media design. As a result, we are likely to underestimate the true effect of the change in testing standards on enrollment in teacher preparation programs. This is for two reasons. First, CIP code 13 includes students not directly affected by the policy change. Second, the increase in testing standards may induce substitution away from teacher preparation programs to other programs within CIP 13. As a robustness check in Section 6.3 we use data from Title II as an alternative (and narrower) measure of enrollment.

IPEDS provides annual data on the number of graduates with more granularity (at the six-digit CIP code level) than on enrollments (at the two-digit CIP code level). Accordingly, our measure of the number of graduates from teaching preparation programs at the institution-level uses these detailed statistics of program completion by major. Measuring graduations at the six-digit CIP code level allows us to focus on the sub-sample of education students who have a major that requires them to pass a teacher preparation entry exam. We

follow the definition of teacher preparation programs recommended by Kraft et al. (2020) and sum the number of graduates across all teacher preparation program majors (undergraduate and master's level) at the institution-level.[11] Unlike enrollment, data on graduations is available annually. Additionally, IPEDS graduation data reflect the academic year. For example, the data point corresponding to 2012 reports graduates from July 2011 until June 2012.

To be included in our sample an institution needed to have at least one year of enrollment data as well as at least one year of data on the number of graduates from teacher preparation programs.[12] The data set for our main results includes 507 universities across 21 states plus DC. Table 2 displays descriptive statistics on the number of education majors and the number of teacher preparation program graduates for the full sample and split by institution-level selectivity. We also report time-varying institution characteristics taken from IPEDS to account for changes in student demographics and institution quality that may impact enrollment or graduations. The variables include counts of full-time-first-time freshman enrollment, SAT and ACT submission rates, the 25^{th} and 75^{th} percentile scores on the SAT and ACT, the percent of students receiving Pell grants and student loans, and the value of Pell grants and student loans conditional on receipt. The average university-year in our sample has 566 students enrolled in the education major and 124 students graduating from a teacher preparation program.

[11]The majors include, among others, 'Education, General', 'Bilingual, Multilingual and Multicultural Education', 'Curriculum and Instruction', 'Special Education and Teaching', 'Teacher Education and Professional Development, Specific Levels and Methods', 'Teaching English or French as a Second or Foreign Language', and 'Education, Other' but excludes categories like "Educational Administration and Supervision".

[12]There are 14 universities that have data on graduates but not enrollment. Results are robust to including these 14 universities.

4 Identification strategy

4.1 Testing standards

To construct a single index that measures the entry testing standards in a state-year, we need to confront three challenges. First, in some states, a test taker who obtains a high enough composite (i.e., combined) score is not bound by the minimum score listed for individual tests by the ETS. Seven states allowed students to pass the PPST as long as their combined scores met a minimum threshold. However, the specific requirements varied: in three states (Maryland, Virginia and Vermont), test takers only needed to meet the combined score requirement, while in the remaining four states (Hawaii, Maine, New Hampshire and Pennsylvania), they also had to reach an alternative minimum score on each test component. This alternative minimum was consistently three points below the official ETS pass score. Due to limited data, we apply a uniform adjustment across all seven states. Specifically, for these seven states we account for the weakening of testing standards caused by the composite score loophole by subtracting three points from the ETS's listed minimum passing score. This approach reflects the reality that test takers in these states could pass with lower individual scores provided they performed better on another component of the test. In 2016, Maine weakened its Praxis Core standards by allowing test takers to achieve a combined score of 468 provided that no single score was more than three points below the minimum passing score. Accordingly, for Maine from 2016 onward, we subtract three points from each test component. In Section 6, we will explore the robustness of our treatment of the composite score loophole.

Second, when ETS replaced the PPST with the Core, it altered the minimum and maximum scores making the comparison of old and new raw passing scores meaningless. To solve this problem, we standardized raw passing scores into z-scores for each component (reading, writing, math) of each test. For each component test, the z-scores are computed as the difference between a state's raw passing score and the national mean score for all test takers

divided by the national standard deviation for that test. The data used to compute these z-scores (national mean test scores and standard deviations for each component of the PPST and Praxis Core) were collected from ETS technical manuals (ETS, 2010, 2015).[13]

Third, we need to convert the three components of each test into a single Test Difficulty Index (TDI). To do this, we took the arithmetic average of the z-scores for each of the three components. Finally, for each state we compute ΔTDI. ΔTDI is equal to the TDI in 2014 (the first year of full adoption of the Praxis Core) minus TDI in 2012 (when the PPST was the only entrance exam). By construction ΔTDI captures the variation in testing standards that is attributable to the fact that all states adopted the scores recommended by the multi-state panel without regard to local concerns.

Table 3 contains the data used to construct the Test Difficulty Index (TDI) for the PPST. Data on the raw passing scores for each state was taken from the ETS's *Praxis Series Passing Score by Test and State* reports and was cross checked with each state's Department of Education. The table reports the raw passing scores and the corresponding z-score for each test component. The PPST z-scores are negative for every state, with many states having minimum passing standards more than one standard deviation below the average score earned by all test takers on one or more test component. As shown in the table, small differences in raw passing scores translate into large differences in z-scores. For instance, for the math test, Connecticut's passing score of 171 translates into a z-score of -1.01 whereas Delaware's passing score of 174 is a z-score of -0.58. Additionally, it is worth noting that states with low z-scores on one component tend to have low z-scores on all components (see, for example, Pennsylvania or Connecticut) with some exceptions (like Maryland or North Carolina).

The second to last column reports the TDI for the PPST. The test difficulty index ranges from a high of -0.25 in Louisiana to a low of -1.52 in Hawaii with an average of -0.83. After

[13]All PPST tests were standardized using national data from the 2010 technical manual since we were unable to locate a 2012 technical manual. For the PPST we standardized using data from those who took the test on paper. We used national data from the 2015 technical manual to standardize the Praxis Core.

the adoption of the Praxis Core, all states had an identical TDI of -0.42.[14]

The final column shows ΔTDI. In one state—Louisiana—TDI decreased with the adoption of the Praxis Core. Thus, teacher preparation entry test standards became easier in Louisiana.[15] In the remaining states entry testing standards increased following the adoption of the Praxis Core, with the magnitude of the increase in the TDI varying substantially across these states. In Alaska, Delaware, North Carolina, Virginia and Wisconsin, testing standards increased by at most 0.12 standard deviations. In contrast, entry requirements were the least binding under the PPST in Connecticut, Hawaii, Mississippi, Nebraska, New Hampshire and Pennsylvania. These six states experienced an increase in testing standards of at least half a standard deviation. Figure 1 shows the Test Difficulty Index for each state year in our sample. The switch from the PPST, where each state had its own passing scores, to the Core, where all states used the same passing scores, is apparent from the figure. All states had a common TDI between 2014 and 2016, after which three states (Maine, Pennsylvania, and South Carolina) weakened their requirements and one state (Connecticut) eliminated the requirement.

Figures 2A and 2B are scatter plots that display the raw correlations between ΔTDI (the change in testing standards between the first full year of the adoption of the Praxis Core and the last year of the PPST) and the percentage change in enrollments and graduations at the state level from 2012 to 2016. In each figure, the best-fit line shows a negative relationship, although it is stronger for graduations than enrollments. This is suggestive evidence that the tightening of testing standards negatively affected the supply of new teachers.

[14]For the Praxis Core the z-scores are -0.23 in math, -0.87 in reading and -0.15 in writing in 2014. As a consequence, the average increase in exam difficulty was driven by the math and writing components. Pass rate data from South Carolina confirms that the math and writing components of the Praxis Core were the most difficult. The three-year pass rate on the reading component was 84.6 percent, while the pass rates on the math and writing components were only 75.3 and 74.8 percent, respectively. See https://ed.sc.gov/policy/education-laws-legislation/office-of-governmental-affairs/requests-from-general-assembly/praxis-core-data/program-admission-assessment-requirement/. In a robustness check we explore the change in each subject test separately.

[15]Given that Louisiana is the only state where entry standards were relaxed, we exclude Louisiana as a robustness check. As shown in A1 our enrollment and graduation are robust to excluding Louisiana from our analysis but including Louisana aids the precision of the enrollment results.

4.2 Enrollments

We first investigate the effect of the change in entry testing standards caused by the replacement of the PPST by the Praxis Core which occurred in 2013-14 on enrollments in education programs. To do this, we estimate the following two-way fixed-effect model:

$$Y_{u,s,t} = \beta_1 TDI_{s,t} + \alpha_{u,s} + T_t + \gamma V_{s,t} + \theta Z_{u,s,t} + \epsilon_{u,s,t} \tag{1}$$

In this regression $Y_{u,s,t}$ is the natural log of fall enrollment in the education program in university u located in state s in year t. Our variable of interest is $TDI_{s,t}$ which is the value of the test difficulty index in state s in year t.[16] Equation (1) includes a set of university fixed effects $\alpha_{u,s}$, and a set of year fixed effects T_t. As mentioned earlier, we only have enrollment data for even-numbered years. Because the Praxis Core was first available in the spring of 2013 and was fully rolled out by fall 2014, we expect enrollment to be affected without a lag. If the change in teacher testing standards is orthogonal to time-varying state-level factors that influence participation in teacher preparation programs, β_1 will be an unbiased estimate of the impact of testing standards on enrollments.

$V_{s,t}$ is a vector of time varying, state-level controls. We control for economic conditions at the state-year level, specifically the unemployment rate and median income taken from the census. Within $V_{s,t}$ we also include a series of state-level education policy indicators (taken from Kraft et al. (2020)) that may affect teacher supply such as teacher accountability reforms, the elimination of teacher tenure, the increase in the probationary period, the weakening of unions, obtaining a federal Race to the Top grant, and adoption of Common Core standards, as well as the year edTPA was required (taken from Chung and Zou (2023)).

[16] As noted earlier, 18 out of the 22 jurisdictions in our sample have the same pass scores for the Praxis Core throughout the sample period (Maine, Pennsylvania, and South Carolina lowered their passing scores and Connecticut eliminated the requirement in 2017.). Equation (1) includes variation from these changes in entry standards and removes Connecticut from the sample after 2016. However, we are concerned that the change in entry standards may be endogenous. As shown in Appendix Table A1, our findings are robust to holding passing scores constant at their 2014 levels. Additionally, these estimates have the advantage of eliminating any biases that may arise when treatment is staggered.

Finally, $Z_{u,s,t}$ is a vector of time-varying institution-level controls including average 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile SAT and ACT scores, whether or not the university is SAT or ACT test-optional, the percent of students receiving a federal Pell grants and student loans, the average dollar amount of federal Pell grants and student loans awarded, and total enrollment. Given that testing standards are set by states, we cluster standard errors at the state level. Additionally, given the small number of clusters, all standard errors are bootstrapped.

We also study the impact of the change in testing standards arising from the replacement of the PPST with the Praxis Core using an event study framework. This specification allows us to conduct a falsification test of our identifying assumption by estimating leads of the impact of the change in testing standards and examining them for evidence of pre-trends. Formally, the event-study regression equation is as follows:

$$Y_{u,s,t} = \alpha_{u,s} + T_t + \sum_{k=-2}^{k=3} \delta_k \Delta TDI_{s,t-k} + \gamma V_{s,t} + \theta Z_{u,s,t} + \epsilon_{u,s,t} \tag{2}$$

For the event study the key variable of interest is ΔTDI. Recall that ΔTDI is equal to TDI in 2014 (the first year of full adoption of the Core) minus TDI in 2012 (when the PPST was the only entrance exam). ΔTDI only reflects the initial change in testing standards brought about by the switch from the PPST to the Praxis Core. It does not include modifications to Praxis Core that occurred in the latter part of the sample. Accordingly, recent concerns about staggered treatment do not apply. Since our enrollment data are biennial, in equation (2) we trace out the anticipatory effects for years 2008 and 2010 as well as the dynamic post treatment effects for years 2014, 2016, and 2018 (the year 2012 is standardized to zero). All other variables are defined as in equation (1). An event-study provides suggestive evidence of the key identifying assumption which is that other factors that influence enrollments in education programs are uncorrelated with ΔTDI. Given that all states adopted a common test score, we expect this assumption to hold. Additionally, this framework allows us to test the assumption that the impact of the change entry standards on enrollment occurs without a lag.

4.3 Graduations

We now turn attention to the impact of testing standards on graduations from teacher preparation programs. As discussed earlier, IPEDS graduation data reflect the academic year. For instance, the data from 2015 captures all graduates from July 2014 to June 2015. While the exact timing varies by state and university, students are usually required to pass the entrance exam (i.e., the PPST or Core) by the end of their sophomore year or the beginning of their junior year if they are undergraduate students. For instance, according to the guidelines for Delaware, all candidates in an undergraduate teacher education program must pass the Praxis Core Academic Skills for Educators tests in reading, writing, and mathematics after completing 60 credit hours and/or prior to admission to upper-level education courses.[17] However, others do not require it until late in their junior year or, in some cases, even their senior year (see, for instance, Mississippi State's requirements atmsstate.edu). Finally, students pursuing two-year master's level teacher training programs are normally required to pass the test prior to admission.

Though the PPST was officially phased out by ETS on September 1, 2014, as shown in Table 1 some states had already transitioned to the Core by the spring of 2013.[18] In these states, students who underwent the test in spring of their sophomore year are expected to graduate in the 2014-15 academic year and should appear in the IPEDS 2015 graduation data. However, if students took the test in the spring of their freshman year, the impact on graduation numbers would only be evident in 2016. It is also worth noting that many states did not adopt the Praxis Core until the fall of 2014, delaying graduation impacts until 2016. As a result, we anticipate the spring of 2015 to be the first period where the effects of entry

[17] We investigated the entry requirements for several teacher preparation programs with an eye to understanding when students were required to take the PPST or the Praxis Core. Many schools require the submission of scores by the end of the sophomore year. See, for instance, the requirements at the University of Vermont, Kean University, and Eastern Kentucky State.

[18] An analysis of data from the Title II National Teacher Preparation Reports (`https://title2.ed.gov/Public/DataTools/Tables.aspx`) confirms that students were taking both the PPST and the Praxis Core in the spring of 2013. Additionally, our enrollment results, which document a significant decline in enrollment in the fall of 2014, are highly suggestive of the Praxis Core being in place in the spring of 2013.

testing become noticeable and for the effect on the number of graduates to increase over time as more and more teacher preparation program students are exposed to this policy change.

To analyze the impact of entry standards on the number of graduates we estimate equation (1) with three modifications. First, we define our outcome variable as the natural log of the number of graduates from an institution's teacher preparation program in spring of that year. Second, all variables are measured annually instead of only in even-numbered years. Finally, to account for the fact changes in entry testing standard impact graduation with a lag, we lag TDI by one year.

To more closely examine the relationship between TDI and graduations, we estimate equation (3), which is a dynamic treatment effect model that allows us to trace the anticipatory and post-treatment effects when the outcome of interest is graduations.

$$Y_{u,s,t} = \alpha_{u,s} + T_t + \sum_{l=-3}^{l=7} \delta_l \Delta TDI_{s,t-l} + \gamma V_{s,t} + \theta Z_{u,s,t} + \epsilon_{u,s,t} \tag{3}$$

For the graduation event study, we use annual data. Recall that ΔTDI is equal to TDI in 2014 (the first year of full adoption of the Core) minus TDI in 2012 (when the PPST was the only entrance exam). As in equation (2), the leave out year is 2012. We can estimate the impact of the policy change on the size of the graduation cohorts for the years 2009 to 2011 when the changes to the entry exam could not have influenced the number of graduates as well as the graduation cohorts for the years 2013 and 2014 when the changes to the entry exam should not have influenced the number of graduates. As discussed earlier, because students take the test two to three years before graduation, we do not expect the change in testing standards—which occurred in 2013/2014—to affect graduations until 2015. An immediate effect of ΔTDI on graduations is potentially a violation of the no anticipation assumption or evidence of omitted variable bias. Additionally, we expect the effect of the change in testing standards to grow over time as more and more entrants into teaching preparation programs are held to the new standard. All other variables are the same as they were for equation (2).

5 Results

5.1 Enrollments of education majors

We begin with an analysis of the impact of changing the entry requirement for teacher preparation programs on enrollment of education majors. Our prediction is that enrollment should decrease as TDI increases since some potential education majors will be unable to meet the higher testing standard. Column 1 of Table 4 presents the results from estimating equation (1) only including university and year fixed effects. A one-standard deviation increase in the test difficulty index is associated with a 16 percent decrease in enrollments in education majors, statistically significant at the 10 percent level. The average change in TDI for the states in our sample was an increase of 0.41 standard deviations. Thus, the change in entry tests from the PPST to the Praxis Core resulted in a 6.6 percent decline ($= 0.41 \times -0.16$) in the number of education majors (see the "Average effect" row in the Table 4.) The average university in our sample has an education enrollment of 566 students. This university therefore experienced a decline in education major enrollments of 37 students per year. In column 2 we add state-level economic controls, state-level education policy controls, and university-level controls and obtain an identical coefficient with improved precision. This suggests that the changes in entry testing standards are orthogonal to state and university-level factors that influence the new teacher pipeline.

Figure 3 presents the event-study coefficients. Consistent with the exogeneity of the adoption of uniform passing scores, we find little evidence of pre-trends. In the years when the PPST was in place, the coefficients are small and imprecisely estimated. However, enrollments immediately decline when the PPST was replaced by the Praxis Core. In terms of magnitudes, an increase in test difficulty of one-standard deviation reduced enrollments by 19 percent in 2014, rising to 22 percent in 2016, and then falling to 17 percent in 2018.[19]

Given our hypothesis that the higher entry testing standards are responsible for the

[19]Coefficient estimates and standard errors from the enrollment event study are shown in tabular format in the first column of Appendix Table A2.

decline in enrollment, we expect the effects to be larger in universities with academically weaker student bodies. To test this conjecture, we use the 2010 IPEDS data on SAT/ACT test scores to classify universities into more selective and less selective sub-samples. Many institutions did not report ACT or SAT scores in IPEDS. We suspect that the vast majority of these schools are open enrollment and therefore impute their ACT or SAT scores to be low. We then classify as more selective any institution with 25^{th} percentile SAT verbal or ACT verbal scores above the in-sample median. All remaining schools are classified as less selective. We estimate equation (1) separately for each sub-sample. The results for less selective universities are displayed in Table 4 columns (3) while column (4) contains the estimates for more selective universities. The point estimates suggest that the effect is concentrated in less selective institutions. A one-standard deviation increase in testing standards is correlated with a 27 percent decline in enrollments among less selective institutions and a seven percent decline in selective institutions. However, these coefficients are imprecisely estimated.

A growing body of research investigates the impact of having a same-race teacher on a wide variety of K-12 student outcomes, including test scores, disciplinary outcomes, social-emotional development ratings, the likelihood of completing high school, and the likelihood of enrolling in college (see, for instance, Morgan and Hu (2023); Lindsay and Hart (2017); Wright et al. (2017); Dee (2005, 2004). Much of the evidence finds positive effects. Given the importance of minority teachers and their limited supply, it worth investigating how the increase in entry testing standards arising from the replacement of the PPST with the Praxis Core affected enrollments of education students of different racial backgrounds.

From 2010 onward, IPEDS reports reliable enrollment data disaggregated by racial categories. To study the impact of entry testing standards on enrollments by race, we construct two mutually exclusive categories, white and non-white, and then estimate equation (1) separately for each category. As shown in columns (5) and (6) of Table 4, the coefficients on the test difficulty index are negative for both groups and statistically indistinguishable from

zero [20].

To further explore the impact of higher entry standards on enrollments, we estimate the model separately for states with growing and shrinking school-age populations. In states with growing school-age populations, institutions may have in place weaker entry standards to increase the supply of potential teachers (i.e., a lower GPA requirement, or a low minimum SAT score, or not failing out students who perform poorly in their first year). In these states, a given increase in test difficulty is likely to have a relatively large impact on education major enrollment because the left tail of the quality distribution has not already been culled. Conversely, in shrinking states, the absence of pressure to produce more teachers may allow institutions' to impose stricter entry standards. Accordingly, we expect the same increase in test difficulty to have a larger impact in states where the school-age population is growing than in states with shrinking school-age populations. We classify a state as growing (shrinking) if its 5-17 year old population increased (decreased) between 2009 and 2018. Among growing states, the school-age population rose an average of 4.8 percent over this period while in shirking states the school-age population declined by 4.5 percent on average. Table 4, column (7) shows the results for shrinking states while column (8) shows the results for growing states. In both instances, the coefficient on testing standards is negative but imprecisely estimated.

5.2 Graduations from teacher preparation programs

We now turn attention to our analysis of the effect of the change in entry testing standards on graduations from teacher preparation programs. A priori, it is possible that an increase in testing standards could reduce enrollments in education majors without affecting graduations. Higher entry standards, by eliminating the left-tail of the quality distribution, may

[20]Given the large disparities in licensing scores by race found in other settings, one might expect the negative effects of higher testing standards to be concentrated among non-white students. There are several reasons why we do not find this pattern. First, the data are noisy, and reporting on race is uneven. Additionally, non-white is a broad category that encompasses many groups, not all of whom are academically disadvantaged (e.g., Asians). Finally, our baseline results are based on students admitted into traditional pathways, where the racial gap in academic preparation may be narrower.

simply remove from the pool those students who would have been unable to graduate. In this case, stricter testing standards are potentially welfare enhancing, since they weed out students who would never have completed teacher preparation programs, saving them time and money.

Table 5 presents our estimates of the impact of the change in testing standards on graduations from teacher preparation programs. As shown in column (1), when we only include university and state fixed effects, a one standard deviation increase in testing standards is associated with a decline in graduations of 20 percent, statistically significant at the 10 percent level. When we include state and university control variables (column (2)), the point estimate increases to 22 percent and is more precisely estimated. TDI increased by 0.41 standard deviations on average in the states in our sample. This implies a reduction of 9.0 percent ($= 0.41 \times -0.22$) in the number of teacher preparation program graduates (see "Average effect" row in Table 5). The average university in our sample has a graduating class of 124 students. In such an institution, graduations would have fallen from 124 to 113 students. Given that there are 507 institutions in our sample this suggests a reduction in the number of new teachers of 5,658 per year across these 22 states. Given that we cannot reject that the the impact on graduations is the same as enrollments, our findings suggest that the increase in entry standards screened out many students who would have otherwise been able to meet exit requirements and graduate from teacher preparation programs.

As before we divide our sample into less and more selective institutions according to their SAT or ACT scores. The results for less selective institutions are shown in column (3) of Table 5, while column (4) displays the results for more selective institutions. The point estimates suggest that a one-standard deviation increase in test difficulty reduces graduations in less selective institutions by 28 percent, while for more selective institutions, graduations fall by 20 percent. However, the estimate for less selective institutions is not statistically significant. We also estimate the model separately for white and non-white students using data from 2012 onward due to poor reporting of graduations by racial category in the earlier

years in the sample. As shown in columns (5) and (6) of Table 5, the increase in testing difficulty reduced graduations for white as well as non-white teacher preparation students by a similar magnitude. Finally, columns (7) and (8) of Table 5 report our findings when we divide the sample into states with shrinking and growing school-age populations, respectively. We find that a one-standard deviation increase in testing standards reduces graduations by 18-22 percent in both groups of states, with the estimate being more precise for shirking states.

Figure 4 presents the event study coefficients from equation (3). Recall that the Praxis Core was first available in the spring of 2013 and that the test is normally taken by students during their sophomore or junior years or prior to admission to a master's program. Thus, the earliest we would expect to see a decline in graduations is 2015. As shown in the figure, the coefficient estimates become economically different from the pre-period in 2015 and reaches statistical significance in 2016. This timing of effects is consistent with our priors. A one-standard deviation increase in TDI reduces graduations by 26 percent in 2016 and peaks at 39 percent in 2019.[21] As shown by the null coefficients in the pre-period years, there is little evidence that other factors that impact graduation rates are correlated with the change in the difficulty of the testing standards.

6 Robustness tests and other implications

6.1 Alternative measures of TDI

Our first robustness check involves re-estimating our two-way fixed effect models on enrollments and graduations using alternative measures of the Test Difficulty Index (TDI). The results for enrollments are shown in panel A of Table 6, while the corresponding estimates for graduations are presented in panel B. Recall that our main measure of test difficulty is

[21] See the second column Appendix Table A2 for the coefficients and standard errors from the graduations event study regression.

the arithmetic average of the z-scores for each of the three tests (math, reading, and writing). In columns (1) through (3) we show the results using each component's z-score as the sole determinant of TDI. For both enrollments and graduations, the coefficient estimate is largest for the math component, suggesting that the math exam is the largest gatekeeper. For graduations, the increase in testing standards for each component has a statistically significant and negative impact. However, the coefficient estimates are all smaller than those for our preferred measure of TDI. This is unsurprising, as some states—such as North Carolina—experienced only modest increases in one component but substantially larger increases in the other two, making any individual test a poor proxy for the overall change in testing standards. Because averaging across components may understate the change in testing difficulty for states where only some components changed substantially, in column (4) we instead use the z-score of the component with the largest change when states switched from the PPST to the Praxis Core. This produces an equal or larger (more negative) coefficient than our averaged measure, reinforcing the finding of a statistically significant negative effect of increased testing standards on both enrollments and graduations.

Finally, our baseline approach constructs the test difficulty index (TDI) accounting for the fact that in seven states there is a composite score loophole. While we believe this is the correct approach, as a robustness test, we assigned each of these seven states the minimum passing scores listed on the ETS's Praxis Series Core by Test and State to construct an alternative test difficulty index. In column (5), we present results using this alternative construction of TDI which ignores the composite score exemption. For enrollments, the results from this exercise are qualitatively similar to our preferred specification. However, the graduation impacts are attenuated and lose statistical significance.

6.2 Placebo tests

A potential threat to our identification strategy is the possibility that state-level changes in teacher preparation program entry test standards are correlated with broader changes in

college enrollments and graduations. For instance, if other determinants of college attendance like tuition or local economic conditions are correlated with the change in teacher preparation program entrance exam difficulty, then we might be misattributing the decline in teacher preparation enrollments and graduations to test difficulty.

To examine this possibility, we conduct three placebo tests that examine the relationship between the change in test difficulty and enrollments and graduations of students who are not required to take this test. The first placebo test is an event study (identical to equation 2) that focuses on the enrollments of all students except education majors. If our experiment is valid, the increase in testing difficulty should not reduce enrollments of students in non-education majors. However, we may see a small increase in non-education enrollments due to students substituting away from education towards other majors because the entry standard became more difficult. Figure 5A shows the coefficients from an event-study of the switch from the PPST to the Praxis Core on non-education major enrollments. While the pre-trends are noisy, the validity of our experiment is not rejected and there is some suggestive evidence that the increase in testing difficulty induced potential education majors to switch to other majors.

Our next two placebo tests examine graduations. The first is an event study (similar to equation 3) that looks at all graduations excluding those from teacher preparation programs. As shown in Figure 5B, we find no evidence that our key result concerning the fall in graduations from teacher preparation programs can be attributed to other factors that influence graduations overall and happen to be correlated with the increase in test difficulty. If anything, graduations from other majors increased in states that saw large increases in entry standards. The second is another event study that looks at graduations of education majors who are not are subject to the entry test requirement (for instance, education administration, school counseling, and instructional media majors). This is our preferred placebo test because it focuses on a subset of students who should also be affected by state-level changes in education policy or funding but are not affected by the entry test. A downward

trend in graduations from other education majors following the switch from the PPST to the Praxis Core would be suggestive of broader changes driving our main results. The results from this exercise are shown in Figure 5C, which shows no strong relationship between the test difficulty index and the number of graduates from other education majors.

6.3 Title II data and other outcomes

Our analysis so far has used data from IPEDS. We have relied on this source because it provides a long pre-period. However, information on teacher preparation pipeline is also available from Title II reports. Both data sets measure enrollments and completions from teacher preparation programs, but there are differences between the two sources.[22] As a validation check, we re-estimate our two-way fixed effect models on enrollments and completions from traditional teacher preparation programs using annual Title II data. Unfortunately, the Department of Education website only makes this data available from 2011.

The results of the enrollment regressions are shown in Panel A of Table 7. As before, we see that increases in testing standards are associated with a reduction in enrollment. In terms of magnitudes, a one-standard deviation increase in testing standards correlates with a 36 percent reduction in enrollments, a larger effect than we observe in the IPEDS data. This corresponds with an average decline in enrollments of 15 percent for states in our sample. These effects are only present in states where the school age population is shrinking and are found for both white and non-white students. Panel B of Table 7 shows the corresponding results for completions. A one-standard deviation increase in testing standards correlates with a statistically significant 30 percent reduction in graduations, an effect similar to what we found using IPEDS.[23] For graduations, the negative effects are concentrated in states with

[22] Title II focus on the enrollment of individuals in teacher preparation programs and whether they complete their programs' requirements while IPEDS focuses on the number of college students by major. See Goldhaber et al. (2017) for more details.

[23] Because the impact of changing testing standards on graduations occurs with a lag, we have enough pre-period data to generate a meaningful event study when graduations are the outcome. As shown in Appendix Figure A2, we find large, negative effects on graduations which attain statistical significance beginning in 2016, consistent with the impact of testing standards on graduations occurring with a lag.

declining school age enrollments and among non-white students. While the effect patterns differ slightly from IPDES, the Title II data are consistent with the view that increases in testing standards reduce graduations and enrollments.

An advantage of Title II data is that it also provides information on enrollment and graduates from alternative teacher preparation pathways. Some states have sizable alternative teacher preparation programs which provide students with a non-education bachelor's degree a pathway to teaching certification. These programs have different entry standards than traditional pathways. It is possible that, in response to rising entry standards, students may substitute into these alternative programs, especially since these programs can be set up and expanded relatively quickly. How this affects the pipeline depends on whether students in alternative programs are subject to passing the ETS entry exams. Many states require students in alternative pathways to meet the state-specific Praxis passing scores if their GPA and/or SAT scores are below a certain threshold.[24] This implies that rising testing standards may still function as an entry barrier for academically marginal students who wish to enter alternative programs. The final columns of Table 7 show the regression results when enrollments and graduations from alternative programs are used as the dependent variable. The coefficient on testing standards is negative and statistically insignificant for both outcomes. We therefore find no evidence that availability of alternative teacher preparation programs lessens the observed decline in the new teacher pipeline.

An important policy question is whether the reduction in teacher preparation program graduates induced by the replacement of the PPST with the Praxis Core in fact reduced the number of new teachers eligible to work. Title II collects state-level information on the number of new teacher licenses awarded each year regardless of type (i.e., traditional licenses awarded to graduates of teacher preparation program, temporary or emergency licenses, and

[24]For instance, alternate route candidate must achieve the Mississippi State Board of Education approved admission criteria of a twenty-one (21) ACT (or SAT equivalent) or achieve a qualifying passing score on the Praxis Core Academic Skills for Educators examination or have verification of a 3.0 overall GPA on a baccalaureate transcript or last 60 hours course credit as established by the State Board of Education prior to non-traditional program admission.

certifications granted to graduates of alternative programs as well as Teach for America).[25] This is the most comprehensive measure of the flow of new teachers into the profession. We estimate a model similar to the graduation event-study regression model (equation 3) at the state-year level using the log of the number of new teacher licenses as the dependent variable (excluding the university-level controls and replacing university fixed effects with state fixed effects) for the years 2010 to 2019 . Figure 6 displays the results. Reassuringly, as with graduations from teacher preparation programs, there is little evidence of differential pre-trends. Additionally, the effect of increased entry testing standards on the number of newly licensed teachers is negative, with the largest declines occurring in 2016 and 2017. This pattern is reassuring, as it aligns with the timing of our graduation results, which begin to show declines in 2015. This is consistent with the fact that there may be a lag between when students graduate from teacher preparation programs and when they meet all licensing requirements.[26]

As additional evidence that testing standards affect the supply of new teachers, we investigated the relationship between changes in testing standards and a proxy for teacher subject area shortages. The U.S. Department of Education collects state-level information on the number of designated teacher shortage areas each year through reports submitted by states to determine teacher eligibility for federal student aid programs.[27] This is one of the few consistent nation-wide sources of information on teacher shortages, though it should be viewed as an imperfect proxy because it reflects the number of shortage categories identified rather than the number of unfilled teaching positions.

To investigate this relationship, we estimate a regression similar to the new teacher

[25]We are grateful to Melissa Lyon and Matthew Kraft for compiling and sharing these data (Kraft and Lyon, 2024).

[26]When we estimate the corresponding two-way fixed effects regression, the coefficient on testing standards is -0.26 and statistically significant at the 10 percent level, indicating a decline in new teacher licenses following an increase in testing standards. In addition, a scatter plot of the change in testing difficulty for each state against the change in the number of new licenses awarded between 2012 and 2016 reveals a negative relationship (see Appendix Figure A3).

[27]See tsa.ed.gov for more information. States may differ in how broadly or narrowly they define shortage areas, but these differences are absorbed by state fixed effects, assuming reporting practices remain stable over time.

licenses analysis, where the dependent variable is the log of the number of teacher shortage areas in a state for the years 2010 to 2020. Figure 7 displays the results of this exercise. As with our other outcomes, there is limited evidence of differential pre-trends. The results suggest that beginning in 2018, increased entry standards are associated with an increase in the number of reported shortage areas. This consistent with the expected timing: it takes time for more demanding entry standards to reduce the flow of new teachers, for these reductions to translate into staffing difficulties, and for states to register those difficulties in their shortage-area designations. We therefore interpret this evidence as broadly supportive of our claim that higher testing standards resulted in economically meaningful reductions in teacher supply, while recognizing that the shortage-area measure is an indirect and imperfect proxy for labor market conditions.

7 Conclusion

In this paper we investigate the impact of a quasi-experimental change in testing standards brought about by the replacement of the PPST with the Praxis Core on the new teacher pipeline. The new test substantially increased entry standards in most states, whereas other states experienced only minor changes in their requirements. Using institution-level data from 21 states plus DC, we find that the average increase in testing standards reduced education major enrollments as well as graduations from teacher preparation programs. Our estimates indicate that an increase in test difficulty of one-standard deviation reduced education major enrollments by 16 percent and teacher preparation program graduations by 22 percent. We also find that the increase in test difficulty resulted in fewer newly licensed teachers and increased staffing difficulties. There is no evidence that the change in entry standards correlates with state-specific trends in education majors or graduates from teacher preparation programs.

Our estimates suggest that the replacement of the PPST with the Praxis Core, which on

average raised entry standards, had an economically large impact on teacher supply, and that it accounts for a sizeable portion of the ongoing decline in graduations from teacher preparation programs in some states. Between 2012 and 2020, the number of teacher preparations program graduates fell by 27 percent in the states in our sample. Our estimates suggest that the change in entry testing standards results in a 9.0 percent decline in graduates for the average state in our sample. The replacement of the PPST with the Praxis Core therefore accounts for 33 percent ($\approx 0.090 \div 0.27$) of the decline in graduates from teacher preparation programs. The three most impacted states in our sample—Hawaii, New Hampshire and Pennsylvania–saw an increase in testing difficulty of 1.10, 0.94 and 0.93 standard deviations, respectively, This implies a reduction in the number of teacher preparation program graduates of 21 percent ($\approx 0.96 \times 0.22$) in these three states. Between 2012 and 2020, the number of teacher preparation program graduates declined by 33 percent in these three states. A back-of-the-envelope calculation therefore suggests that 66 percent ($= 0.22 \div 0.33$) of this decline can be explained by the high minimum passing scores that were put in place with the adoption of the Praxis Core. Accordingly, current teacher shortages in many states may, at least in part, be an unintended consequence of the replacement of the PPST with the Praxis Core.

An important question is whether this increase in testing standards improved teacher quality, in terms of student learning outcomes. As discussed by Angrist and Guryan (2008), an increase in entry testing standards could increase or decrease student learning. On the one hand, if higher entry standards induce an improvement in the average effectiveness of the teaching pool, student learning may improve. On the other hand, student learning may suffer since higher entry standards reduce teacher supply, which in turn may force school districts to hire fewer or less qualified teachers. Additionally if the production function for effective teaching is multidimensional (for instance, it may require teacher academic ability as well as classroom management skills), high testing standards may select along academic ability at the expense of classroom management, which could reduce teacher effectiveness (Dinerstein

and Opper, 2022). Given that the first cohort of teachers to be affected graduated in the spring of 2015 and entered the classroom that fall, only a small share of the stock of teachers were exposed to the higher entry standard. Accordingly, ours is not an ideal setting in which to investigate the impact of this policy change on student learning.[28]

While we are agnostic about the effects of higher testing standards on teacher quality, we believe that our findings can explain why state-level policy makers have, in response to perceived teacher shortages, recently pushed for the lowering of Praxis passing scores, or even the wholesale elimination of entry requirements like the Praxis Core. According to a 2021 study published by the National Council on Teacher Quality (Putman and Walsh, 2021), many states have lowered or removed academic requirements for entry into teacher preparation programs. West Virginia now permits teacher training programs to waive test requirements or to admit students based on a cohort average instead of requiring individual students to pass the test. Mississippi, where testing entry standards increased by 0.75 standard deviations, reduced the pass score for the math portion of the Praxis Core by 20 points in 2019. In 2022 and 2023 respectively, Louisiana and Nebraska, eliminated the requirement for teachers to use the Praxis Core to demonstrate basic skills.[29] Finally, Pennsylvania, where testing standards increased by almost one standard deviation, introduced a temporary waiver of the Praxis Core requirement in 2022 and permanently eliminated the Praxis Core requirement in 2025. Whether these policy changes reverse the decline in teacher preparation program enrollments and alleviate the shortage of new teachers is a topic for future research.

[28]Recent evidence suggests that Covid-era waivers of teacher licensure requirements had no effect on teacher effectiveness (Backes and Goldhaber, 2023; Bacher-Hicks et al., 2023).

[29]See Nicholls State University (2022) for Louisiana. For Nebraska see ne.gov (Accessed: November 29, 2023)

References

Angrist, J. and Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, 27(5):483–503.

Bacher-Hicks, A., Baloch, S., Chi, O., and Tichnor-Wagner, A. (2023). Teacher licensure and workforce quality: Insights from emergency licenses in massachusetts. Wheelock Educational Policy Center. Available at: `http://wheelockpolicycenter.org`.

Backes, B. and Goldhaber, D. (2023). The relationship between pandemic-era teacher licensure waivers and teacher demographics, retention, and effectiveness in new jersey. working paper no. 286-0623. *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*.

Barrios, J. M. (2022). Occupational licensing and accountant quality: Evidence from the 150-hour rule. *Journal of Accounting Research*, 60(1):3–43.

Blair, P. Q. and Chung, B. W. (2019). How much of barrier to entry is occupational licensing? *British Journal of Industrial Relations*, 57(4):919–943.

Buddin, R. and Zamarro, G. (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics*, 66(2):103–115.

Chen, B., Dougherty, S., Goldhaber, D., Holden, K., and Theobald, R. (2023). CTE Teacher Licensure and Long-Term Student Outcomes. *Education Finance and Policy*, 18(2):253–276.

Chung, B. W. and Zou, J. (2023). Teacher Licensing, Teacher Supply, and Student Achievement: Nationwide Implementation of edTPA. EdWorkingPapers 21-440, Annenberg Institute at Brown University.

Chung, B. W. and Zou, J. (2025). Occupational licensing in us public schools: Nationwide implementation of teacher performance assessment. *Journal of Public Economics*, 244:105328.

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.

Decker, P., Mayer, D., and Glazerman, S. (2012). The Effects of Teach For America on Students: Findings from a National Evaluation. Institute for Research on Poverty, Discussion Paper no. 1285-04.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1):195–210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 95(2):158–165.

Dinerstein, M. and Opper, I. (2022). Screening with multitasking. NBER Working Paper 30310, National Bureau of Economic Research.

ETS (2010). *Praxis™ Technical Manual*. Educational Testing Service. Revised December 2010 edition.

ETS (2013). Praxis PPST and CORE comparison chart. Accessed: 2023-09-13.

ETS (2015). *Praxis™ Technical Manual*. Educational Testing Service. Revised 2015 edition.

Gittleman, M., Klee, M. A., and Kleiner, M. M. (2018). Analyzing the labor market outcomes of occupational licensing. *Industrial Relations*, 57(1):57–100.

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.

Goldhaber, D., Gratz, T., and Theobald, R. (2017). What's in a teacher test? Assessing the relationship between teacher licensure test scores and student STEM achievement and course-taking. *Economics of Education Review*, 61:112–129.

Goldhaber, D. and Hansen, M. (2010). Race, Gender, and Teacher Testing: How Informative a Tool Is Teacher Licensure Testing? *American Educational Research Journal*, 47(1):218–251.

Guasch, J. L. and Weiss, A. (1981). Self-selection in the labor market. *The American Economic Review*, 71(3):275–284.

Jacob, B. A. and Lefgren, L. (2004). Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *Review of Economics and Statistics*, 86(1):226–244.

Kane, T. J., Rockoff, J. E., and Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6):615–631.

Kleiner, M. (2013). *Stages of Occupational Regulation: Analysis of Case Studies*. W.E. Upjohn Institute, Kalamazoo, MI.

Kleiner, M. M. (2006). *Licensing Occupations: Ensuring Quality or Restricting Competition?* W.E. Upjohn Institute, Kalamazoo, MI.

Kleiner, M. M. and Krueger, A. B. (2010). The prevalence and effects of occupational licensing. *British Journal of Industrial Relations*, 48(4):676–687.

Kleiner, M. M. and Krueger, A. B. (2013). Analyzing the extent and influence of occupational licensing on the labor market. *Journal of Labor Economics*, 31(2):S73–S202.

Kraft, M. A., Brunner, E. J., Dougherty, S. M., and Schwegman, D. J. (2020). Teacher accountability reforms and the supply and quality of new teachers. *Journal of Public Economics*, 188:104212.

Kraft, M. A. and Lyon, M. A. (2024). The rise and fall of the teaching profession: Prestige, interest, preparation, and satisfaction over the last half century. *American Educational Research Journal*, 61(6):1192–1236.

Larsen, B., Ju, Z., Kapor, A., and Yu, C. (2020). The effect of occupational licensing stringency on the teacher quality distribution. NBER Working Paper 28158, National Bureau of Economic Research.

Law, M. T. and Kim, S. (2005). Specialization and regulation: The rise of professionals and the emergence of occupational licensing regulation. *Journal of Economic History*, 65(3):723–756.

Law, M. T. and Marks, M. S. (2009). Effects of occupational licensing laws on minorities: Evidence from the Progressive Era. *Journal of Law and Economics*, 52(2):351–366.

Law, M. T. and Marks, M. S. (2017). The labor market effects of occupational licensing laws in nursing. *Industrial Relations*, 56(4):640–661.

Lindsay, C. and Hart, C. (2017). Teacher-student race match and student disciplinary outcomes in north carolina. *Educational Evaluation and Policy Analysis*, 39(3):485–510.

Morgan, P. L. and Hu, E. H. (2023). Fixed effect estimates of student-teacher racial or ethnic matching in us elementary schools. *Early Childhood Research Quarterly*, 63:98–112.

Nicholls State University (2022). Praxis Core education program requirement removed by BESE. `https://www.nicholls.edu/news/2022/09/02/praxis-Core-education-program-requirement-removed-by-bese/`. Accessed: November 27, 2023.

Orellana, A. and Winters, M. A. (2025). Licensure tests and teacher supply. *Journal of Human Resources*.

Pagliero, M. (2010). Licensing exam difficulty and entry salaries in the US market for lawyers. *British Journal of Industrial Relations*, 48(4):726–739.

Pagliero, M. (2013). The impact of potential labor supply on licensing exam difficulty. *Labour Economics*, 25:141–152.

Putman, H. and Walsh, K. (2021). *Teacher Preparation Policy*. State of the States 2021. National Council on Teacher Quality, Washington, DC.

Sass, T. R. (2015). Licensure and Worker Quality: A Comparison of Alternative Routes to Teaching. *Journal of Law and Economics*, 58(1):1–35.

Schmitt, J. and deCourcy, K. (2022). The pandemic has exacerbated a long-standing national shortage of teachers. Economic Policy Institute Report No. 254745.

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.

Thornton, R. J. and Timmons, E. J. (2013). Licensing one of the world's oldest professions: Massage. *Journal of Law and Economics*, 56(2):371–388.

Will, M. (2022). Fewer people are getting teacher degrees. Prep programs sound the alarm. *Education Week*, 41(28):14.

Wright, A., Gottfried, M. A., and Le, V.-N. (2017). A kindergarten teacher like me: The role of student-teacher race in social-emotional development. *American Educational Research Journal*, 54:78S–101S.

Tables

Table 1: State Descriptive Information

State	PPST	Core Adoption Date	Adopted Pass Scores	Panel Participant	In Sample	Shrinking State
AK	yes	Fall 2014	yes	yes	yes	no
AR	yes	Fall 2013	yes	yes	yes	no
CT	yes	Fall 2014	yes†	yes	yes	yes
DE	yes	Spring 2014	yes	yes	yes	no
DC	yes	Fall 2013	yes	yes	yes	yes
HI	yes	Fall 2013	yes	yes	yes	no
IN	yes	n/a	n/a	no	no	n/a
IA	no	n/a	n/a	yes	no	n/a
KY	no	Fall 2014	yes	yes	no	n/a
LA	yes	Fall 2014	yes	yes	yes	yes
ME	yes	Fall 2013	yes†	yes	yes	yes
MD	yes	Fall 2014	yes	yes	yes	no
MS	yes	Fall 2013	yes	yes	yes	yes
NE	yes	Fall 2014	yes	yes	yes	no
NV	yes	Fall 2013	yes	no	yes	no
NH	yes	Fall 2013	yes	yes	yes	yes
NJ	yes	Fall 2013	yes	yes	yes	yes
NC	yes	Fall 2013	yes	yes	yes	no
ND	yes	Fall 2014	no*	yes	no	n/a
OK	yes	+	+	no	no	n/a
OR	yes	Spring 2014	yes	no	yes	no
PA	yes	Fall 2014	yes†	no	yes	yes
RI	no	n/a	n/a	yes	no	n/a
SC	yes	Fall 2013	yes†	yes	yes	no
TN	yes	Fall 2013**	yes	yes	no	n/a
VT	yes	Fall 2013	yes	yes	yes	yes
VA	yes	Fall 2013	yes	no	yes	no
WA	no	Fall 2014	no	no	no	n/a
WV	yes	Fall 2013	yes	yes	yes	yes
WI	yes	Fall 2014	yes	yes	yes	yes

Only that states used the PPST or Praxis Core or participated in the multi-state standard setting panel are included in this table. Shrinking states are defined as states that saw their school-age (5-17) population shrink from 2008 to 2018.

† On September 1, 2016, Connecticut eliminated its requirement that students pass the Praxis Core, although they must still take the test. Maine added a composite score adjustment in 2016. On September 1, 2016, Pennsylvania lowered its math passing score. In June 2017, South Carolina lowered its math and writing passing scores (retroactive to September 2016).

* North Dakota adopted the recommended pass score for the math and reading test but chose a lower score on the writing test.

+ Oklahoma adopted the Praxis Core in Fall 2014 but allowed candidates to substitute the Oklahoma General Education Test.

** ETS reports that TN requires the Praxis Core while the TN Department of Ed states "Praxis Core academic tests may be required for admission to an educator preparation program, but may not be required for licensure."

Table 2: Summary Statistics

	All	More Selective	Less Selective
Panel A Outcomes:			
Education fall enrollments	566	720	398
	(769)	(762)	(74)
White education fall enrollments	390	454	195
	(527)	(556)	(401)
Non-white education fall enrollments	170	266	203
	(333)	(439)	(447)
Teacher preparation graduates	124	165	73
	(169)	(188)	(122)
White teacher preparation graduates	88	115	44
	(123)	(139)	(83)
Non-white teacher preparation graduates	29	50	29
	(70)	(99)	(63)
Panel B Time-varying Controls:			
Total freshmen enrollment	892	1259	458
	(1135)	(1379)	(545)
SAT submission rate	63.5	63.0	61.0
	(30.1)	(31.3)	(27.2)
ACT submission Rate	41.3	43.1	41.3
	(29.0)	(30.0)	(26.6)
SAT math 25^{th} percentile score	472	494	446
	(56.8)	(56.1)	(45.3)
SAT verbal 25^{th} percentile score	470	493	446
	(55)	(53.4)	(45.5)
ACT composite 25^{th} percentile score	19.5	20.7	18.1
	(2.9)	(2.9)	(2.1)
Pell grant (dollar amount, thousands)	4232	4220	4340
	(520)	(471)	(598)
Pell grant (% student)	40.7	31.3	52.1
	(18.6)	(12.7)	(18.2)
Loan (dollar amount, thousands)	7252	7544	6993
	(1766)	(1725)	(1741)
Loan (% student)	67.3	64.2	70.5
	(18.0)	(17.0)	(18.1)
Number of universities	507	267	240

Notes: All university data comes Integrated Postsecondary Education Data System. School age population change is calculated using the American Community Survey. Less selective universities are those with scores below the median or those that did not report test scores. More selective universities are defined as those with SAT/ACT scores above the in-sample median as of 2008. The table presents mean values, with standard deviations shown in parentheses underneath.

Table 3: Descriptive Information on Testing Standards

	Pre-Professional Skills Test (PPST)							
	Math		Reading		Writing			
State	Raw Score	Z-Score	Raw Score	Z-Score	Raw Score	Z-Score	TDI	ΔTDI
AK	173	-0.72	175	-0.38	174	-0.38	-0.50	0.08
AR	171	-1.01	172	-0.93	173	-0.64	-0.86	0.44
CT	171	-1.01	172	-0.93	171	-1.15	-1.03	0.62
DC	174	-0.58	172	-0.93	171	-1.15	-0.89	0.47
DE	174	-0.58	175	-0.38	173	-0.64	-0.53	0.12
HI*	170	-1.16	169	-1.47	168	-1.92	-1.52	1.10
LA	175	-0.43	176	-0.20	175	-0.13	-0.25	-0.17
MD*	174	-0.58	174	-0.56	170	-1.41	-0.85	0.43
ME*	172	-0.87	173	-0.75	172	-0.90	-0.84	0.42
MS	170	-1.16	172	-1.29	173	-0.90	-1.16	0.75
NC	173	-0.72	176	-0.20	173	-0.64	-0.52	0.11
NE	171	-1.01	170	-1.29	172	-0.90	-1.07	0.65
NH*	169	-1.30	171	-1.11	169	-1.67	-1.36	0.94
NJ	174	-0.58	175	-0.38	173	-0.64	-0.53	0.12
NV	172	-0.87	174	-0.56	172	-0.90	-0.78	0.36
OR	175	-0.43	174	-0.56	171	-1.15	-0.72	0.30
PA*	170	-1.16	169	-1.47	170	-1.41	-1.35	0.93
SC	172	-0.87	175	-0.38	173	-0.64	-0.63	0.21
VA*	175	-0.43	175	-0.38	173	-0.64	-0.48	0.06
VT*	172	-0.87	174	-0.56	171	-1.15	-0.86	0.44
WI	173	-0.72	175	-0.38	174	-0.38	-0.50	0.08
WV	172	-0.87	174	-0.56	172	-0.90	-0.78	0.36
Average	173.2	-0.82	174.2	-0.71	172.9	-0.92	-0.83	0.41

Notes: "Raw Score" is the minimum score needed to pass each subject of the Pre-Professional Skills Test. Data on passing scores came from the Education Testing Service cross-checked with each state's Department of Education. For the seven states with composite score loopholes (denoted by *) we subtracted three points from the pass score listed by ETS. The z-scores are computed as the difference between a state's raw passing score and the national mean score for all test takers divided by the national standard deviation for that test. TDI is the simple average of the three z-scores shown in columns 3, 5, and 7. ΔTDI is the change in Test Difficulty Index from the last year of the PPST to the first year of the Praxis Core.

Table 4: Composite Enrollments: Total Enrollments

	(1) All	(2) All	(3) Less Selective	(4) More Selective	(5) White	(6) Non White	(7) Shrinking State	(8) Growing State
TDI	-0.16+	-0.16*	-0.27	-0.07	-0.31	-0.10	-0.16	-0.12
	(0.10)	(0.08)	(0.20)	(0.09)	(0.23)	(0.14)	(0.10)	(0.37)
Observations	2,896	2,896	1,313	1,583	2,896	2,896	1,567	1,329
State Level Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
University Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean enrollment	566	566	398	720	548	596	387	691
Num. universities	507	507	240	267	505	505	240	267
Average effect	-0.07	-0.07	-0.11	-0.03	-0.13	-0.04	-0.07	0.05

Notes: Each column represents a separate regression. The dependent variable in each regression is the log of the number of fall education enrollments. All regressions include university and year fixed effects. State-level controls include the unemployment rate, median income, and education policy indicators. University controls include ACT and SAT test scores, enrollment, and financial aid variables. More selective universities are defined as those with SAT/ACT scores above the in-sample median as of 2010. In contrast, less selective universities are those with scores below the median or those that did not report test scores. Columns 5 and 6 are estimated using data from 2012 onward. Growing states are defined as states that saw their school-age (5-17) population increase from 2008 to 2018. Average effect is an estimate of the change in enrollments for the average state in the sample (the point estimate on TDI multiplied by the average change in TDI for that sample). Bootstrapped standard errors clustered at the state level are displayed in parentheses. ** $p<0.01$, * $p<0.05$, + $p<0.10$.

Table 5: Teacher Preparation Program Graduates

	(1) All	(2) All	(3) Less Selective	(4) More Selective	(5) White	(6) Non–White	(7) Shrinking State	(8) Growing State
TDI	-0.20+	-0.22*	-0.28	-0.20**	-0.24**	-0.19*	-0.22**	-0.18
	(0.11)	(0.10)	(0.20)	(0.07)	(0.05)	(0.08)	(0.06)	(0.12)
Observations	5,748	5,748	2,617	3,131	4,325	4,309	3,098	2,650
State level controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
University controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean graduates	124	124	72	161	90	28	121	128
Num. universities	507	507	240	267	503	503	272	235
Average effect	-0.08	-0.09	-0.11	-0.08	-0.10	-0.08	-0.09	-0.07

Notes: Each column represents a separate regression. The dependent variable in each regression is the log of the number of teacher preparation program graduates in a university-year. All regressions include university and year fixed effects. State-level controls include the unemployment rate, median income, and education policy indicators. University controls include ACT and SAT test scores, enrollment, and financial aid variables. More selective universities are defined as those with SAT/ACT scores above the in-sample median as of 2010. In contrast, less selective universities are those with scores below the median or those that did not report test scores. Columns 5 and 6 are estimated using data from 2012 onward. Growing states are defined as states that saw their school-age (5-17) population increase from 2008 to 2018. Average effect is an estimate of the change in graduates for the average state in the sample (the point estimate on TDI multiplied by the average change in TDI for that sample). Bootstrapped standard errors clustered at the state level are displayed in parentheses. ** $p<0.01$, * $p<0.05$, + $p<0.10$.

Table 6: Robustness Tests Using Alternative Measures of the Test Difficulty Index

Panel A: Education Enrollments

	(1) Math Test	(2) Reading Test	(3) Writing Test	(4) Binding Test	(5) ETS Min Scores
TDI	-0.17	-0.12+	-0.11+	-0.16*	-0.17
	(0.12)	(0.07)	(0.06)	(0.08)	(0.08)
Observations	2,880	2,880	2,880	2,880	2,880
Mean enrollment	566	566	566	566	566
Number of universities	507	507	507	507	507

Panel B: Graduations

	(1) Math Test	(2) Reading Test	(3) Writing Test	(4) Binding Test	(5) ETS Min Scores
TDI	-0.21**	-0.16+	-0.18*	-0.29**	-0.08
	(0.08)	(0.08)	(0.09)	(0.09)	(0.11)
Observations	5,748	5,748	5,748	5,748	5,748
Mean graduates	124	124	124	124	124
Number of universities	507	507	507	507	507

Notes: Each column represents a separate regression. The dependent variable in Panel A is the log of the number of fall education enrollments and in Panel B is log of the number of graduates from teacher preparation programs. All regressions include university fixed effects, year fixed effects, state economic controls , state education policy controls, and university-level controls as specified in equation (1). Bootstrapped standard errors clustered at the state level are displayed in parentheses. ** p<0.01, * p<0.05, + p<0.10.

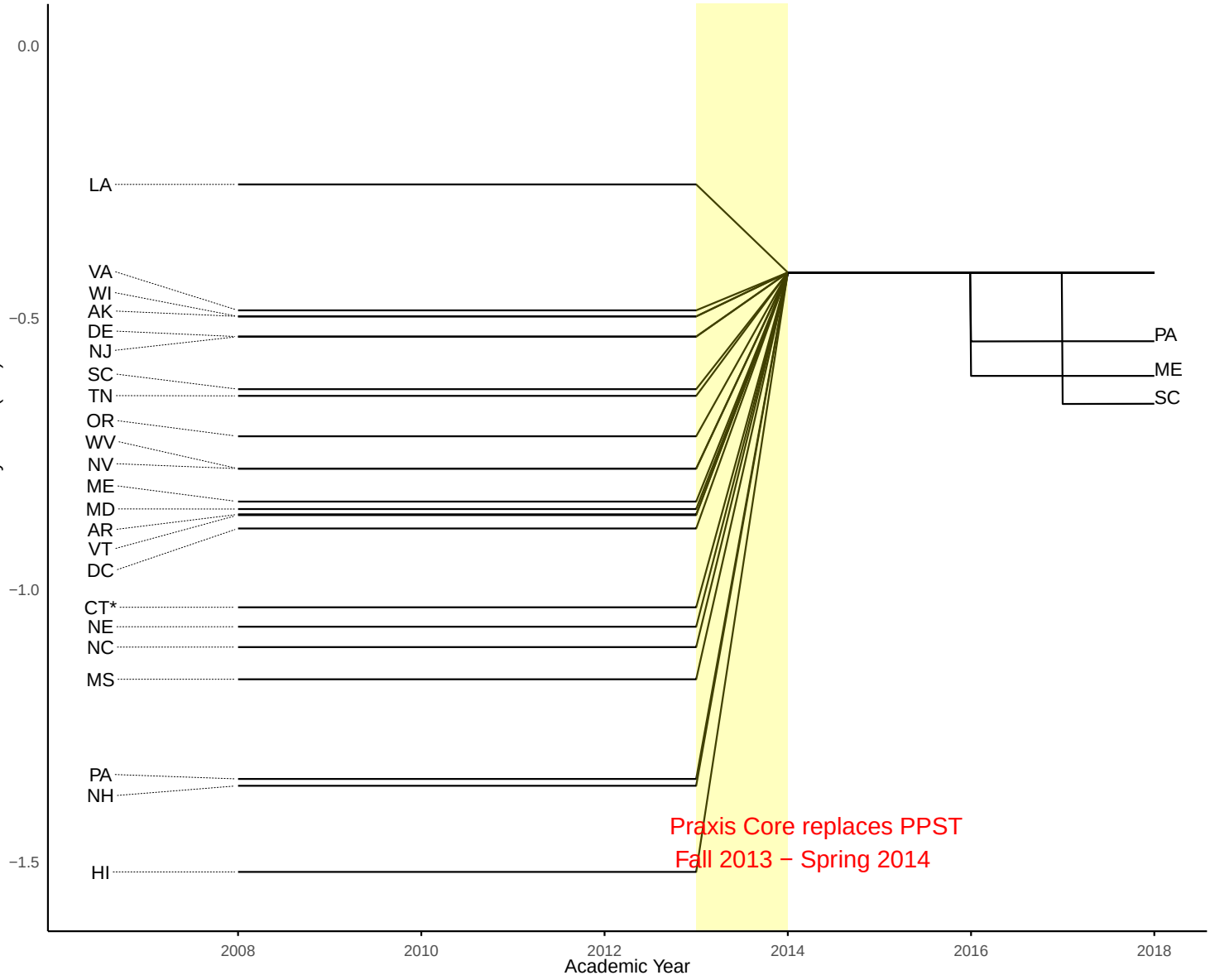
Table 7: Title II Enrollments and Graduations

	Traditional Programs					Alternative
	(1) All	(2) Shrinking State	(3) Growing State	(4) White	(5) Non–White	(6) All
Panel A: Enrollments						
TDI	-0.36**	-0.45**	-0.02	-0.32*	-0.38*	-0.09
	(0.14)	(0.09)	(0.28)	(0.14)	(0.18)	(0.35)
Observations	4,738	2,562	2,176	4,738	4,715	1,912
Mean enrollment	231	223	241	169	61	109
Number of programs	847	464	383	847	847	364
Average effect	-0.15	-0.18	-0.10	-0.13	-0.16	-0.04
Panel B: Graduations						
TDI	-0.30**	-0.32*	-0.13	-0.12	-0.36*	-0.35
	(0.11)	(0.13)	(0.14)	(0.11)	(0.15)	(0.46)
Observations	4,742	2,562	2,180	4,742	4,742	1,915
Mean graduates	85	85	84	11	74	43
Number of programs	847	464	383	847	847	364
Average effect	-0.12	-0.13	-0.05	-0.05	-0.15	-0.14

Notes: Each column represents a separate regression. In panel A the dependent variable is the log of the number of teacher preparation program enrollees as the dependent variable while in panel B it is the log of the number of teacher preparation program graduates. All regressions include program fixed effects, year fixed effects, and state economic and education policy controls. Growing states are defined as states that saw their school-age (5–17) population grow from 2008 to 2018. Average effect is an estimate of the change in enrollment for the average state in the sample (the point estimate on TDI multiplied by the average change in TDI for that sample). Bootstrapped standard errors clustered at the state level are in parentheses. ** $p < 0.01$, * $p < 0.05$.

Figures

Figure 1: Test Difficulty Index



Notes: This figure shows the value the Test Difficulty Index for each state in our sample from 2008 until 2018. The Test Difficulty Index is an equally weighted sum of the z-scores for the three components of the either the Praxis Core or the PPST. Time is defined as of January 1st in year t. In 2016 CT eliminated the requirements that students pass the Praxis Core. In the fall of, 2016 Pennsylvania changed their Praxis Core math cutoff from 150 to 142 and Maine changed its requirements to allow test takers to achieve a combined score of 468 with no single score lower than three points lower than the minimum passing score. In June 2017, South Carolina changed its Praxis core math cutoff from 150 to 142 and writing from 162 to 158, with a retroactive effective date of September 1, 2016.

Figure 2A: The Relationship Between the Change in Test Difficulty Index and the Change in Fall Education Enrollments

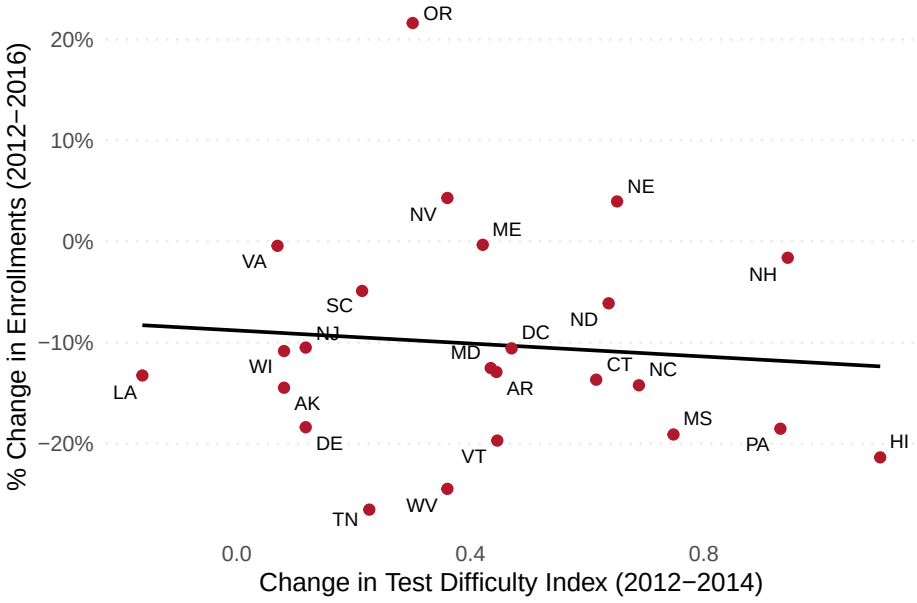
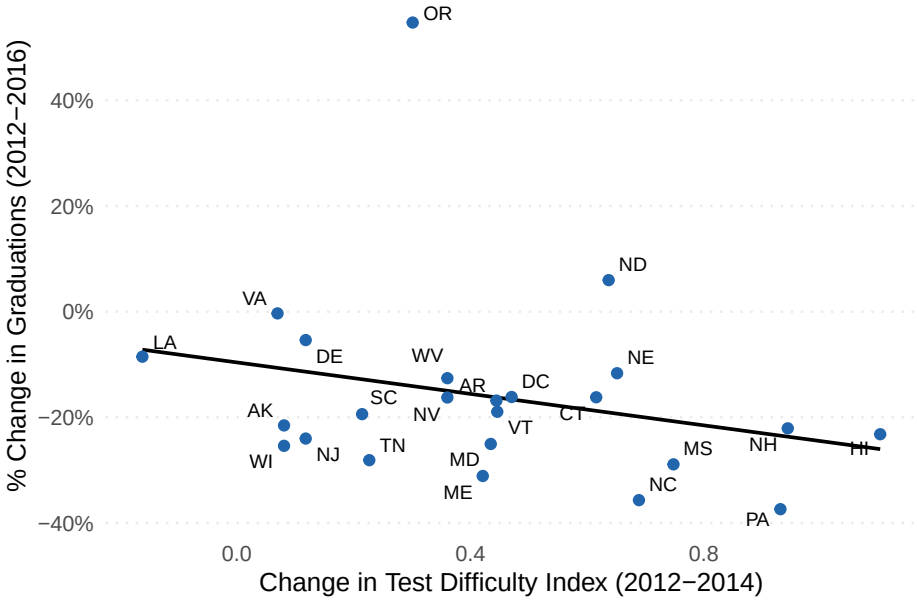
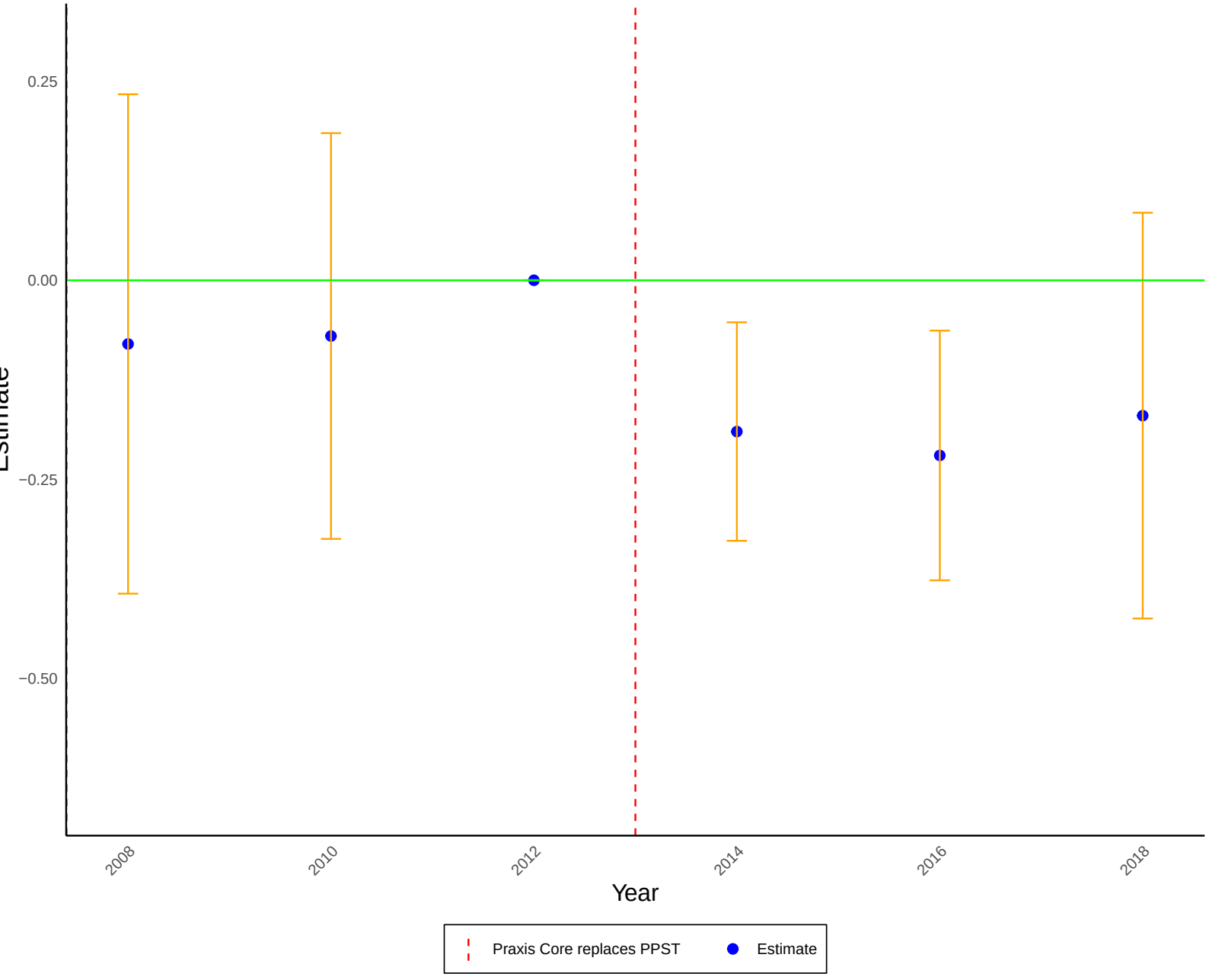


Figure 2B: The Relationship Between the Change in Test Difficulty Index and the Change in Teacher Preparation Program Graduates



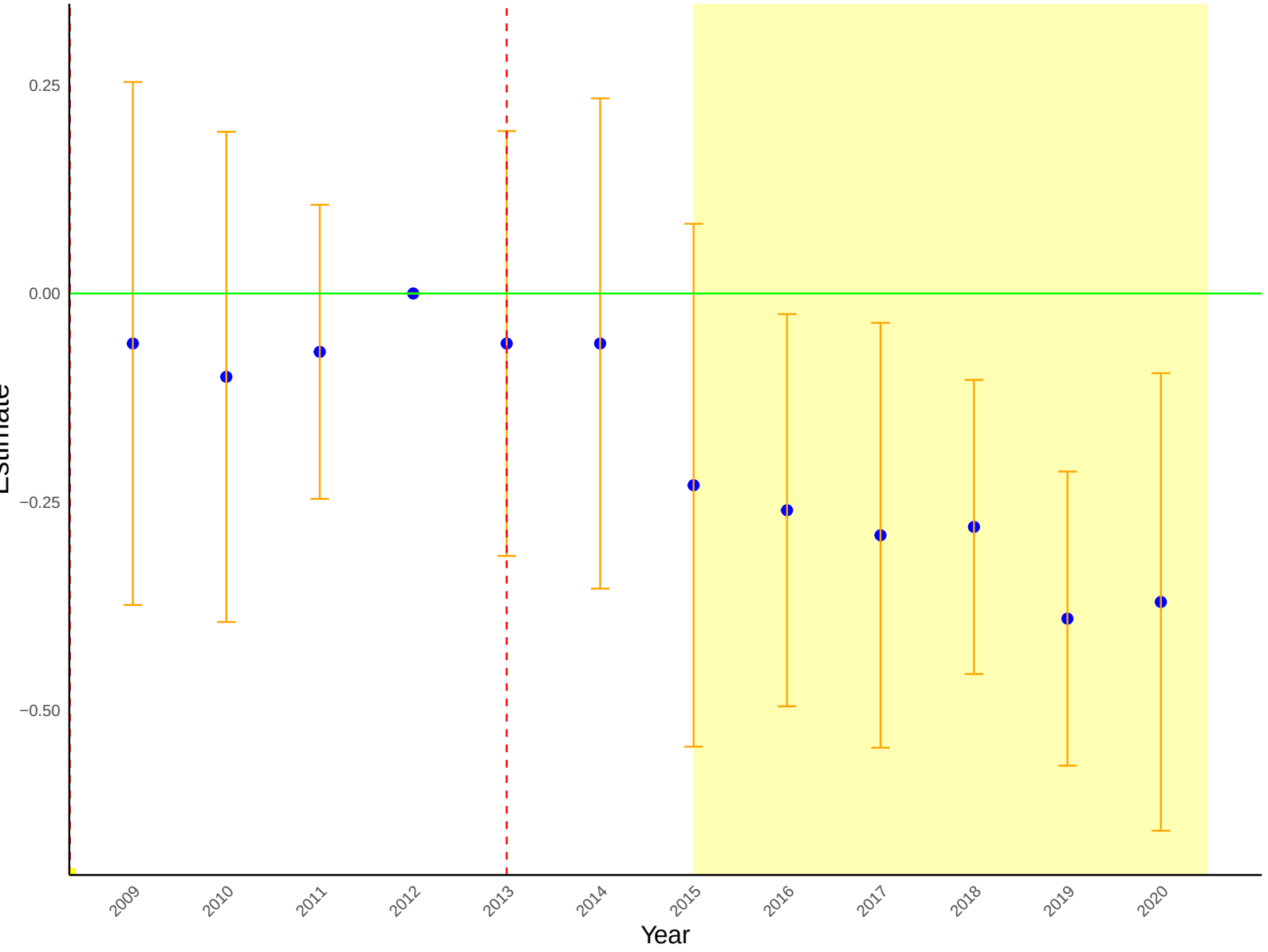
Notes: For both figures, university-level counts are aggregated to the state level.

Figure 3: Education Major Enrollments Event Study



Notes: This graph depicts the estimated coefficients from Equation (2) with log fall education major enrollments as the dependent variable. The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Figure 4: Teacher Preparation Program Graduations Event Study



Notes: This graph depicts the estimated coefficients from Equation (3) with log teacher preparation program graduates as the dependent variable. The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Figure 5A: Enrollment in Non Education Majors

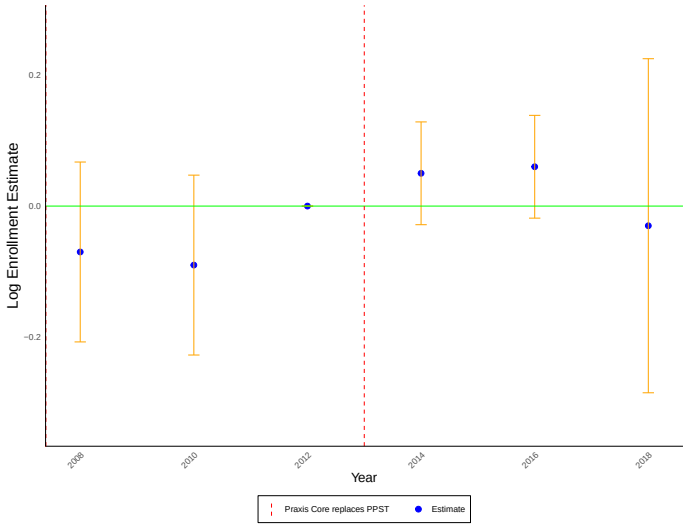


Figure 5B: Non Education Graduations

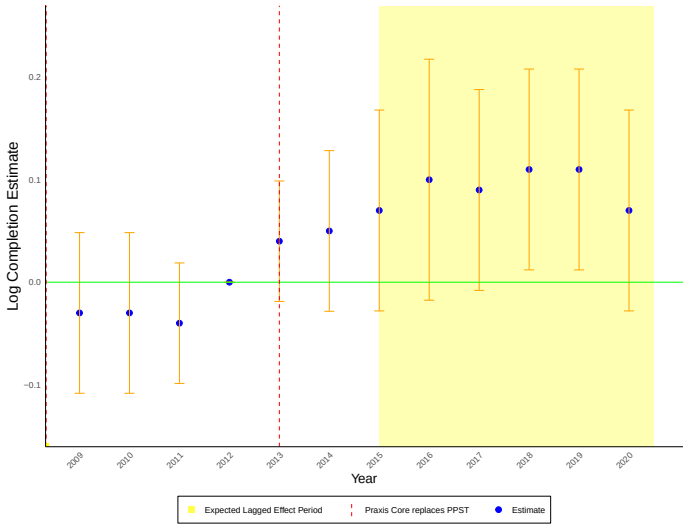
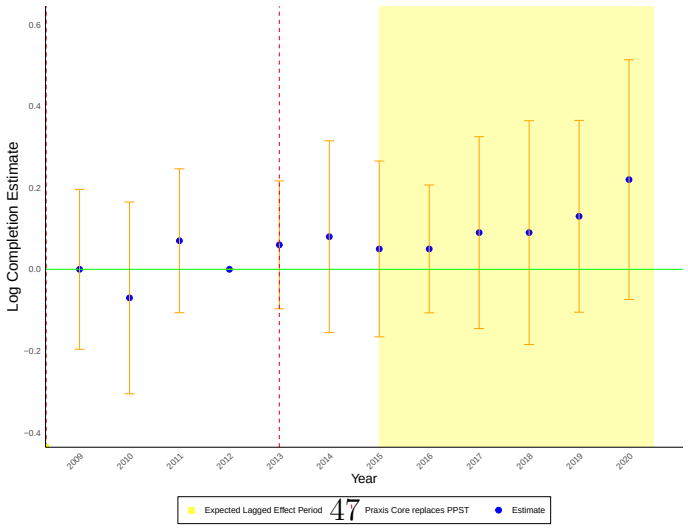
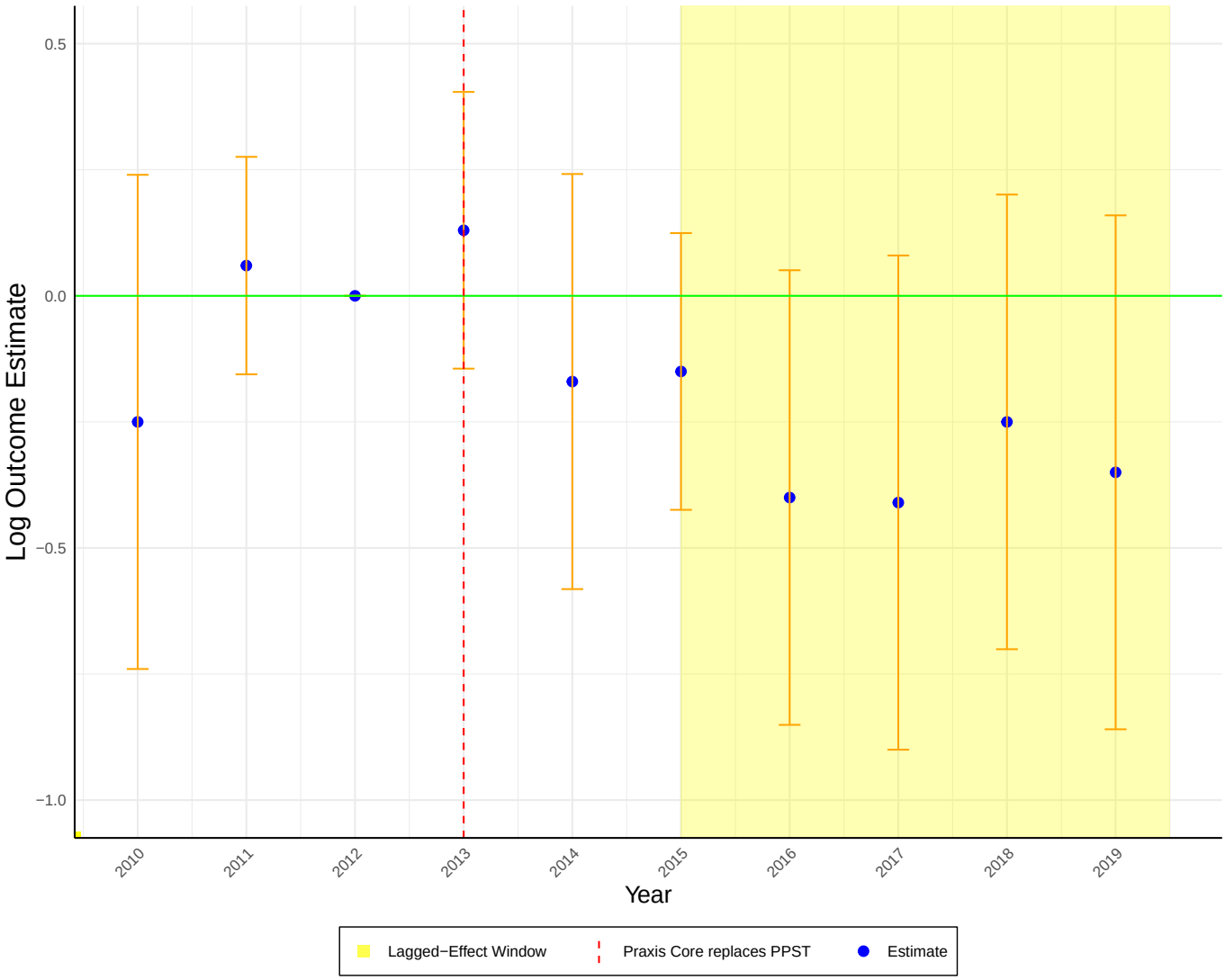


Figure 5C: Other Education Graduations



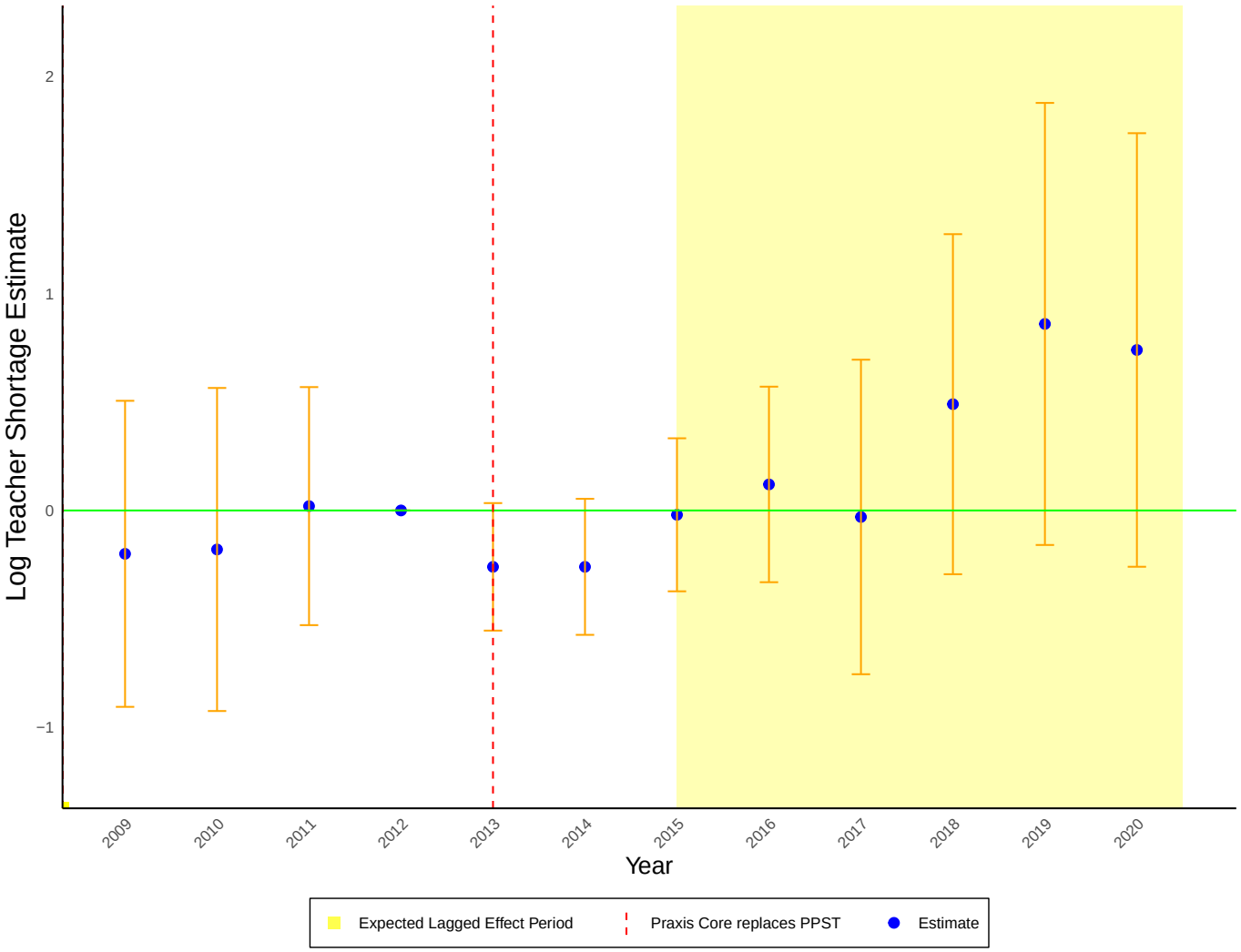
Notes: Figure 5A shows the estimated coefficients from Equation (2), while Figures 5B and 5C show the estimated coefficients from Equation (3). The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Figure 6: Log New Teacher License Estimate



Notes: This graph depicts the estimated coefficients from an event study where the dependent variable is log of new teacher licenses at the state-year level which includes all new teacher licenses awarded each year regardless of type. The model includes state and year fixed effects, state economic controls, and education policy controls. The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Figure 7: Teacher Shortage Area Event Study



Notes: This graph shows the estimated coefficients from an event study where the dependent variable is log of the number of teacher shortage areas the state-year level. The model includes state and year fixed effects, state economic controls, and education policy controls. The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Appendix Figure A1 (continued)

PRAXIS PPST Reading and *PRAXIS* Core Academic Skills for Educators: Reading Comparison Chart

The **Pre-Professional Skills Test in Reading** measures the ability to understand, analyze, and evaluate written texts. Varying in difficulty, the reading material is drawn from a variety of subject areas and real–life situations that educated adults are likely to encounter. Passages are drawn from both print and electronic media, such as newspapers, magazines, journals, nonfiction books, novels, online articles, and visual representations. Each passage is followed by questions that are based on its content and that relate to a variety of reading skills. All questions can be answered by using information contained within the passage; no question requires outside knowledge of the content.

The **Core Academic Skills for Educators Test in Reading** measures academic skills in Reading needed to prepare successfully for a career in education. All skills assessed have been identified as needed for college and career readiness, in alignment with the Common Core State Standards for Reading. In particular, there is an emphasis on skills that are critical to learning and achievement in teacher preparation programs. These skills include the ability to understand, analyze, and evaluate texts of different kinds. Varying in difficulty, the reading material on the test is drawn from a variety of subject areas and real-life situations that educated adults are likely to encounter. Passages are drawn from both print and electronic media, such as newspapers, magazines, journals, nonfiction books, novels, online articles, and visual representations. Each passage is followed by questions that are based on its content and that relate to reading skills. All questions can be answered by using information contained within the passage; no question requires outside knowledge of the content.

	PPST Reading (5710/0710)	**Core Reading (5712)**
Number of questions	46 multiple-choice questions (5710) 40 multiple-choice questions (0710)	56 selected-response questions
Time allotted	75 minutes (5710) 60 minutes (0710)	85 minutes
Delivery Method	Computer-delivered (5710) Paper-delivered (0710)	Computer-delivered
Content Categories	I. Literal Comprehension II. Critical and Inferential Comprehension	I. Key Ideas and Details II. Craft, Structure, and Language Skills III. Integration of Knowledge and Ideas
Types of passages	• Long passages (approx. 200 words) with 4-7 questions • Short passages (approx. 100 words) with 2 or 3 questions • Brief statements with 1 question	• Paired passages totaling approx. 200 words with 4-7 questions • Long passages (approx. 200 words) with 4-7 questions • Short passages (approx. 100 words) with 2 or 3 questions • Brief statements with 1 question
Types of questions	Single-selection multiple-choice questions	Selected-response questions, including • Single-selection multiple-choice questions • Multiple-selection multiple-choice questions • Select-in-passage questions
Scoring scale	150 - 190	100 - 200

Copyright © 2013 by Educational Testing Service. All rights reserved. ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). PRAXIS and THE PRAXIS SERIES are trademarks of ETS.

Appendix Figure A1 (continued)

PRAXIS PPST Writing and _PRAXIS_ Core Academic Skills for Educators: Writing Comparison Chart

The **Pre-Professional Skills Test in Writing** assesses the ability to use grammar and language appropriately and the ability to communicate effectively in writing; these abilities are essential to a well-educated adult in a professional role. The multiple-choice section is designed to measure examinees' ability to use standard written English correctly and effectively. The essay section assesses examinees' ability to write effectively in a limited period of time. Experienced teachers read and evaluate each essay holistically (that is, with a single score for overall quality) under carefully controlled conditions designed to ensure fair and reliable scoring.

The **Core Academic Skills for Educators Test in Writing** measures academic skills in writing needed to prepare successfully for a career in education. All skills assessed have been identified as needed for college and career readiness, in alignment with the Common Core State Standards for Writing. The selected-response section is designed to measure examinees' ability to use standard written English correctly and effectively. The two essays assess examinees' ability to write effectively in a limited period of time. The Argumentative essay topic invites examinees to draw from personal experience, observation, or reading to support a position with specific reasons and examples. The Informative/Explanatory essay topic asks examinees to extract information from two provided sources to identify important concerns related to an issue. Experienced teachers read and evaluate each essay holistically (i.e., with a single score for overall quality) under carefully controlled conditions designed to ensure fair and reliable scoring.

	PPST Writing (5720/0720)	**Core Writing (5712)**
Number of questions	44 multiple-choice questions (5720) 40 multiple-choice questions (0720) 1 constructed-response question (both 5710 and 0710)	40 selected-response questions 2 constructed response questions
Time allotted	68 minutes (5720) – 38 minutes for the multiple-choice section and 30 minutes for the constructed-response section 60 minutes (0720) – Two 30-minute sections	100 minutes (40 minutes for the selected-response section and two 30-minute constructed-response sections)
Delivery Method	Computer-delivered (5720) Paper-delivered (0720)	Computer-delivered
Content Categories	I. Grammatical Relationships II. Structural Relationships III. Word Choice and Mechanics IV. Essay	I. Text Types, Purposes, and Production II. Language and Research Skills for Writing
Types of questions	Single-selection multiple-choice questions: • Usage • Sentence Correction Constructed-response question: • Argumentative Essay	Selected-response questions: • Usage • Sentence Correction • Revision in Context • Research Skills Constructed-response questions: • Argumentative Essay • Informative/Explanatory Essay
Scoring scale	150 - 190	100 - 200

Copyright © 2013 by Educational Testing Service. All rights reserved. ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). PRAXIS and THE PRAXIS SERIES are trademarks of ETS.

Appendix Figure A1 (continued)

PRAXIS PPST Mathematics and _PRAXIS_ Core Academic Skills for Educators: Mathematics Comparison Chart

The **Pre-Professional Skills Test in Mathematics** measures those mathematical skills and concepts that an educated adult might need. It focuses on the key concepts of mathematics and on the ability to solve problems and to reason in a quantitative context. Many of the problems require the integration of multiple skills to achieve a solution. Computation is held to a minimum, and few technical words are used. Terms such as area, perimeter, ratio, integer, factor, and prime number are used because it is assumed that these are commonly encountered in the mathematics all examinees have studied. Figures are drawn as accurately as possible and lie in a plane unless otherwise noted.

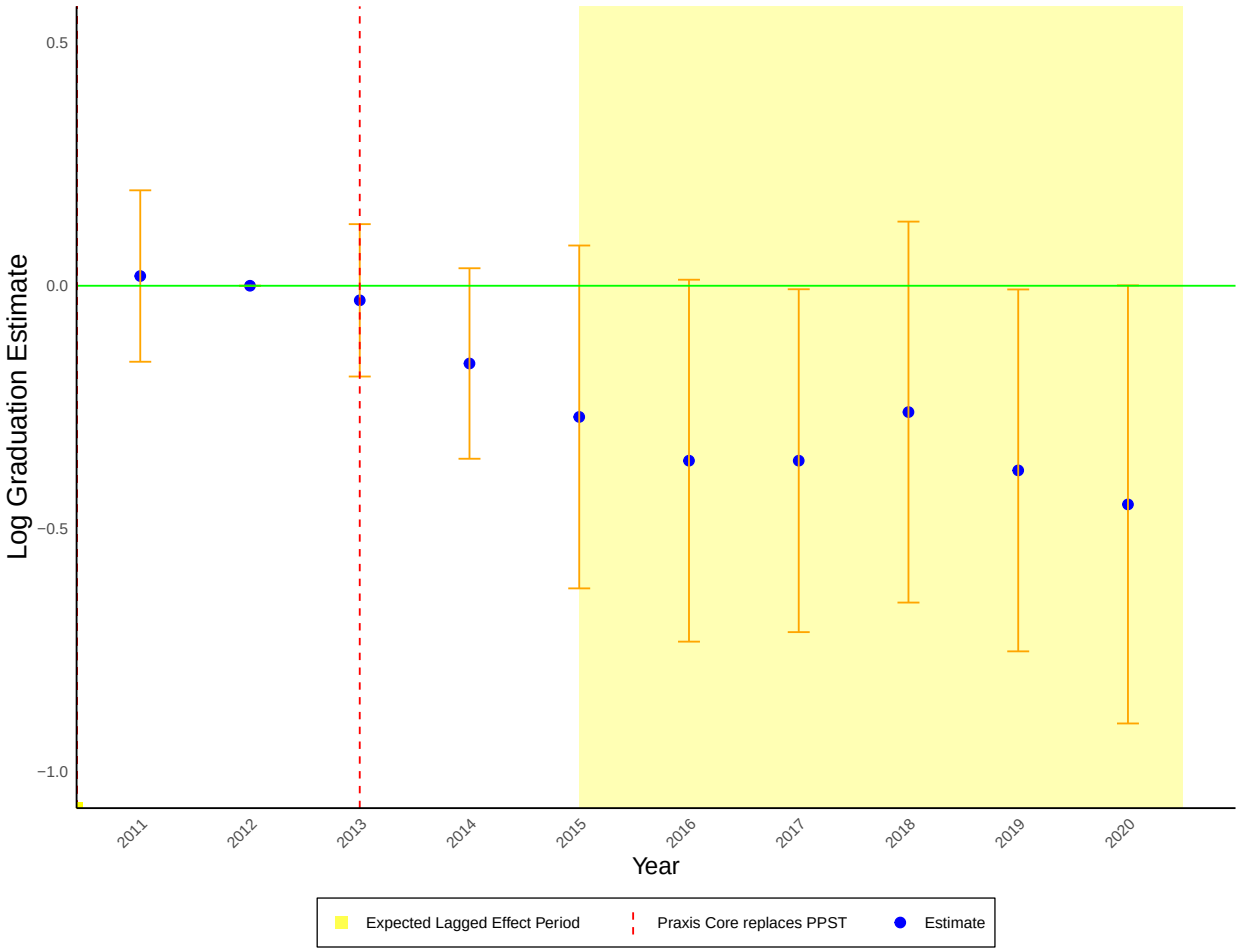
The **Core Academic Skills for Educators Test in Mathematics** measures academic skills in mathematics needed to prepare successfully for a career in education. All skills assessed have been identified as needed for college and career readiness, in alignment with the Common Core State Standards for Mathematics. Focus is on key concepts of mathematics and the ability to solve problems and to reason in a quantitative context. Many of the problems require the integration of multiple skills to achieve a solution.

In Number and Quantity, the understanding of order among integers, representation of a number in more than one way, place value, properties of whole numbers, equivalent computational procedures, ratio, proportion, and percent are emphasized. Algebra assesses the ability to handle equations and inequalities, recognition of various ways to solve a problem, relationship between verbal and symbolic expressions, and graphs. Functions questions test the knowledge of basic function definitions and the relationship between the domain and range of any given functions. Geometry assesses the understanding and application of the characteristics and properties of geometric shapes, the Pythagorean theorem, transformation, and use of symmetry to analyze mathematical situations. Knowledge of basic U.S. customary and metric systems of measurement is assumed. Statistics and Probability assesses the ability to read and interpret visual display of quantitative information, understand the correspondence between data and graph, make inferences from a given data display, determine mean, median, and mode, and assign a probability to an outcome.

	PPST Mathematics (5730/0730)	**Core Mathematics (5732)**
Number of questions	46 multiple-choice questions (5730) 40 multiple-choice questions (0730)	56 selected-response questions
Time allotted	75 minutes (5730) 60 minutes (0730)	85 minutes
Delivery Method	Computer-delivered (5730) Paper-delivered (0730)	Computer-delivered
Content Categories	I. Number and Operations II. Algebra III. Geometry and Measurement IV. Data Analysis and Probability	I. Number and Quantity II. Algebra and Functions III. Geometry IV. Statistics and Probability
Types of questions	Single-selection multiple-choice questions	Selected-response and numeric entry questions, including • Single-selection multiple-choice questions • Multiple-selection multiple-choice questions • Numeric entry questions
Calculator	Calculator not allowed	On-screen four-function calculator provided
Scoring scale	150 - 190	100 - 200

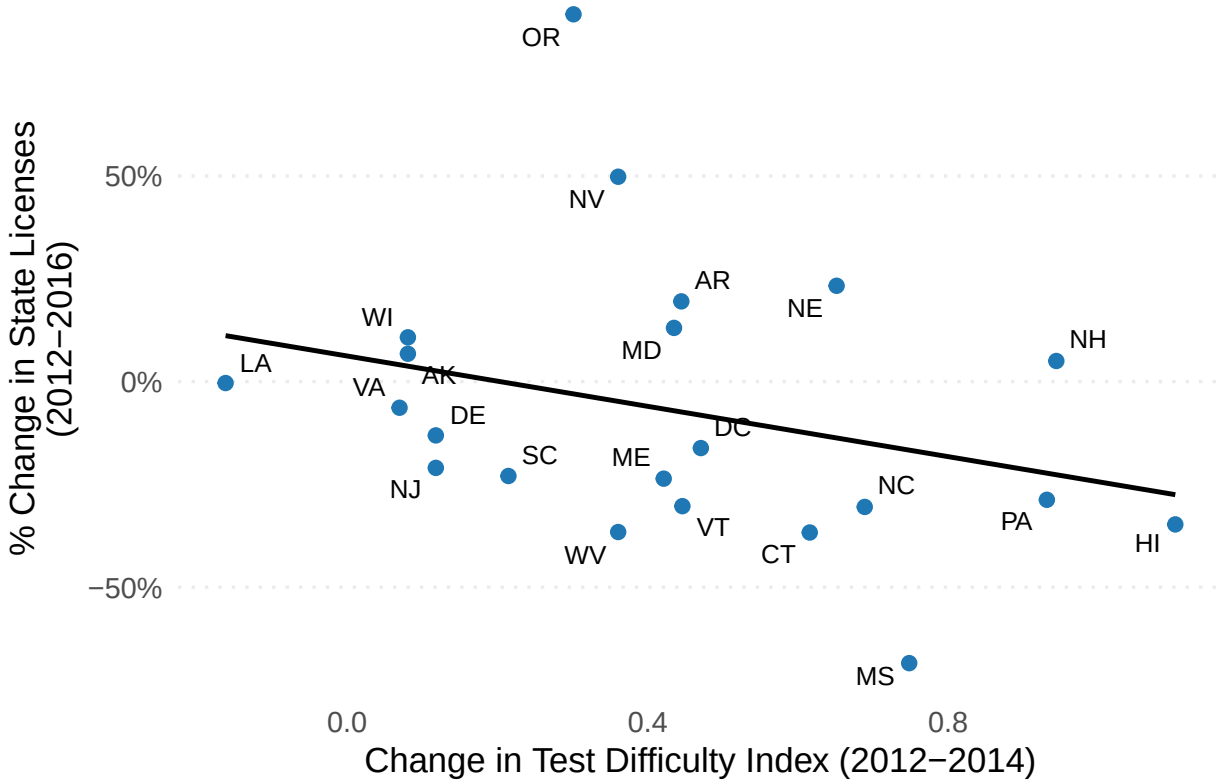
Copyright © 2013 by Educational Testing Service. All rights reserved. ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). PRAXIS and THE PRAXIS SERIES are trademarks of ETS.

Figure A2: Title II Graduations Event Study



Notes: This figure shows the coefficients from an event study where the dependent variable is the log of the number of teacher preparation program completions (measured using Title II data). The lines denote the 95% confidence intervals from a model with bootstrapped standard clustered at the state level.

Figure A3: Relationship Between the Change in Test Difficulty Index and the Change in New Teacher Licenses



Notes: This figure plots the percentage change in TDI between 2012 and 2014 against the percentage change in the number of new teacher licenses between 2012 and 2016 at the state level.

Table A1: Robustness Tests Using Alternative Samples

	(1) Main Sample	(2) Include ND	(3) Include TN	(4) Exclude LA	(5) Fixes ME, SC, PA
Panel A: Enrollments					
TDI	-0.16*	-0.14+	-0.14+	-0.15	-0.15*
	(0.08)	(0.08)	(0.08)	(0.10)	(0.07)
Observations	2,896	2,967	3,127	2,773	2,751
Mean enrollments	566	566	561	563	580
Number of universities	507	520	550	486	507
Panel B: Graduations					
TDI	-0.22*	-0.21*	-0.22*	-0.24**	-0.22*
	(0.10)	(0.11)	(0.09)	(0.08)	(0.10)
Observations	5,748	5,892	6,203	5,504	5,619
Mean graduates	124	124	123	123	125
Number of universities	507	520	550	486	507

Notes: Each column represents a separate regression. The dependent variable in regressions under Panel A is the log of the number of fall education enrollments and in Panel B is the log of the number of teacher preparation program graduates. Column (1) replicates the main results. Column (2) adds North Dakota, which did not adopt all recommended passing scores for the Praxis Core. Column (3) adds Tennessee (ETS and the state department of education disagree on whether the Core is required). Column (4) limits sample to states where TDI increased (Louisiana's TDI decreased). Column (5) holds passing scores fixed at their 2014 levels for Maine, South Carolina and Pennsylvania. Bootstrapped standard errors clustered at the state level are displayed in parentheses. All regression include university fixed effects, year fixed effects, state economic controls, state education policy controls, and university-level controls as specified in equation (1). Bootstrapped standard errors clustered at the state level are shown in parentheses. ** $p<0.01$, * $p<0.05$, + $p<0.10$.

Table A2: Event Study Coefficients

	Enrollments	Graduations
Coefficient: 2008	-0.08	-
	(0.16)	
Coefficient: 2009		-0.06
		(0.16)
Coefficient: 2010	-0.07	-0.10
	(0.13)	(0.15)
Coefficient: 2011		-0.07
		(0.09)
Leave Out Year: 2012		
Coefficient: 2013		-0.06
		(0.13)
Coefficient: 2014	-0.19**	-0.06
	(0.07)	(0.15)
Coefficient: 2015		-0.23
		(0.16)
Coefficient: 2016	-0.22**	-0.26*
	(0.08)	(0.12)
Coefficient: 2017		-0.29*
		(0.13)
Coefficient: 2018	-0.17	-0.28**
	(0.13)	(0.09)
Coefficient: 2019		-0.39**
		(0.09)
Coefficient: 2020		-0.37**
		(0.14)
Observations	2,896	5,784
Mean dependent variable	566	124

Notes: This table presents the event study coefficients from equation (3). The dependent variable in column (1) is the log of the number of education major fall enrollments. The dependent variable in column (2) is the log of the number of teacher preparation program graduates in a university-year. All regressions include university fixed effects, year fixed effects, state economic controls, state education policy controls, and university-level controls as specified in equation (1) Bootstrapped standard errors clustered at the state level are displayed in parentheses. ** p<0.01, * p<0.05