STATE OF CONNECTICUT
STATE BOARD OF EDUCATION

TO: Clerk of the Senate
Clerk of the House of Representatives

FROM: Charlene M. Russell-Tucker, Commissioner of Education

DATE: February 22, 2024

SUBJECT: Content Area Mastery and Educator Certification

In accordance with Section 384 of Public Act 21-2ss, enclosed is a study of a multiple measures approach to demonstrating content-area mastery for the purposes of Connecticut General Statutes Section 10-145f. This study reviews current assessment requirements for educator certification, analyzes candidate first-time and best attempt pass rates, describes supports for candidate access to and use of free-retakes, and introduces multiple pathways to demonstrate content-area mastery.

If you have any questions please contact Laura Stefon, Chief of Staff and Legislative Liaison, at 860-713-6493.

Cc: Education Committee
Legislative Library
Office of Legislative Research
State Library

Enclosure

Study of Content Area Licensure Assessments for Teacher Certification

CSDE

CONNECTICUT STATE DEPARTMENT OF EDUCATION

February 2024

Contents

Background

Unless an individual is already certified in another state that is recognized by Connecticut under enhanced reciprocity, all other individuals looking to become newly certified to teach in Connecticut public schools must successfully complete an approved educator preparation program (EPP) *and* attain a minimum passing score in the appropriate content area standardized assessment(s) as required in Section 10-145f of the Connecticut General Statutes. The primary purpose of these assessments is to ensure **minimum content knowledge** in the content area for individuals entering the teaching profession. Given the academic and cognitive rigor expected of our students by the CT Core Standards in many content areas (see Smarter Balanced sample items in Grade 5 Math and Grade 6 English Language Arts for illustrative purposes), it is essential that all educators including those teaching elementary and middle school have adequate content knowledge so they can explain complex concepts in different ways so all students will learn.

In Connecticut, the most widely used content area assessment is the Praxis II which is developed by the Educational Testing Service (ETS). Praxis II is not a single test but a wide range of assessments. Each unique content area license requires different tests, some requiring one, and others, like elementary, requiring multiple.

To determine the recommended cut/passing scores for all Praxis II exams, ETS conducts multi-state standard-setting activities using psychometrically recognized procedures. The recommended cut scores differentiate the "just-qualified" candidate from the "not-quite qualified" candidate in each content area. Connecticut's passing scores align with the ETS recommended cut-scores for the "just-qualified" candidate.

Legislation

In the 2021 legislative session, the Connecticut General Assembly passed Public Act 21-2ss. Section 384 of this Act required the Connecticut State Department of Education (CSDE) *"to conduct a study of a multiple measures approach to demonstrating content-area mastery for the purposes of section 10-145f of the general statutes. Such study shall include, but not be limited to, a review of current assessment requirements for educator certification, candidate first-time pass rates, best attempt pass rates, candidate access to and use of free-retake policy, and alternative multiple measure pathways to demonstrate content-area mastery for certification"* (p. 602).

Methodology

The CSDE used data from the Educational Testing Service (ETS) and merged it with the CSDE's educator certification and employment data to create a new Educator Preparation Program (EPP) dashboard on EdSight. This dashboard provides a range of metrics including the total number of candidates, number of completers, number of completers certified within one year, pass rates on licensure examinations, and employment in Connecticut public schools. The licensure examination pass rates from 2018-19 and 2021-22 were examined for this study.

The CSDE also engaged Boston University's Wheelock Educational Policy Center (WEPC) to examine the relationship between the state's teacher licensure exam requirements and the composition of its teacher workforce. Specifically, WEPC focused on the following question: **What is the effect of failing the first attempt of a required licensure exam on the likelihood of becoming a teacher in Connecticut?** WEPC also provided the CSDE with information about multiple measure approaches being explored by other states for the demonstration of minimum content area knowledge.

Data on test scores from 1995 to 2021 as well as teacher certification and assignment data from 2002 to 2020 were examined by WEPC. In total, the study sample consisted of about 85,000 individuals with Praxis II test scores; about 51,000 of them also have teacher employment records within the state. To estimate the causal effect of failing the licensure test on becoming a teacher, researchers used a **regression discontinuity design** to exploit the similar attributes of candidates right above and below the passing cut-score for each test and examine the differences in their trajectories into or away from the profession.

As stated previously, Praxis II is not a single test but a wide range of assessments. Each unique content area license requires different tests, some requiring one, and others, like elementary, requiring multiple. Due to the nature of the available data, measuring distinct impacts of individual tests was not feasible. The research team was also unable to examine variations in results based on prospective teachers' race/ethnicity or gender.

Results

Overall test pass rates for program completers from Connecticut EPPs for 2018-19 (pre-pandemic) and 2021-22 (most recent available) are presented below (Table 1). The data are disaggregated by race/ethnicity.

Table 1: Overall test pass rates for program completers from Connecticut EPPs

	2018-19				2021-22			
Race / Ethnicity	# of Completers Attempting a Test	# Tests Attempted	Tests Passed in First Attempt	Tests Passed in Best Attempt	# of Completers Attempting a Test	# Tests Attempted	Tests Passed in First Attempt	Tests Passed in Best Attempt
Candidates of Color	218	623	66.0%	93.1%	214	542	59.0%	86.3%
White Candidates	1,021	2,864	77.3%	97.7%	913	2,493	75.3%	95.0%
Not Available	98	268	82.8%	99.3%	121	277	75.5%	90.6%
Total	**1,337**	**3,755**	**75.8%**	**97.1%**	**1,248**	**3,312**	**72.7%**	**93.2%**

These data show the following:

- Around 1,200-1,300 individuals complete CT EPPs annually and they take over 3,000 tests.
- Prior to the pandemic, approximately three quarters of all licensure tests taken, and around two thirds of all licensure tests taken by candidates of color were passed on their first attempt. Over 90% of tests taken were passed after one or more subsequent attempts.
- First and best attempt pass rates have declined for all groups after the pandemic, though more than 85% among all groups continue to pass after one or more attempts (i.e., best attempt).

In light of the slightly lower pass rates on first attempt for candidates of color, the CSDE examined first, and best attempt pass rates for candidates of color across select institutions (Table 2). The data illustrate considerable variation among EPPs in test pass rates for candidates of color, especially with respect to the first attempt. At most institutions, over 50% of tests taken by candidates of color were passed in their first attempt, and this increased to over 80% when considering all subsequent attempts. At two of

the new EPPs i.e., CREC and Relay GSE, the first attempt pass rates were below 40%, and increased to over 60% when considering all subsequent attempts.

Table 2: 2021-22 Pass Rates for Candidates of Color for Select Institution

Institution	# of Completers (# of Tests Taken)	First Attempt Pass Rate	Best Attempt Pass Rate
Alternate Route to Cert	20 (26)	88.5%	96.2%
CREC	12 (60)	31.7%	66.7%
Central CSU	21 (51)	72.5%	86.3%
Eastern CSU	13 (19)	57.9%	100.0%
Relay GSE (2020-21)*	72 (253)	35.6%	62.8%
Southern CSU	47 (143)	55.9%	79.7%
University of Bridgeport	17 (45)	51.1%	100.0%
University of Connecticut	30 (62)	71.0%	98.4%
University of Hartford	11 (32)	59.4%	90.6%

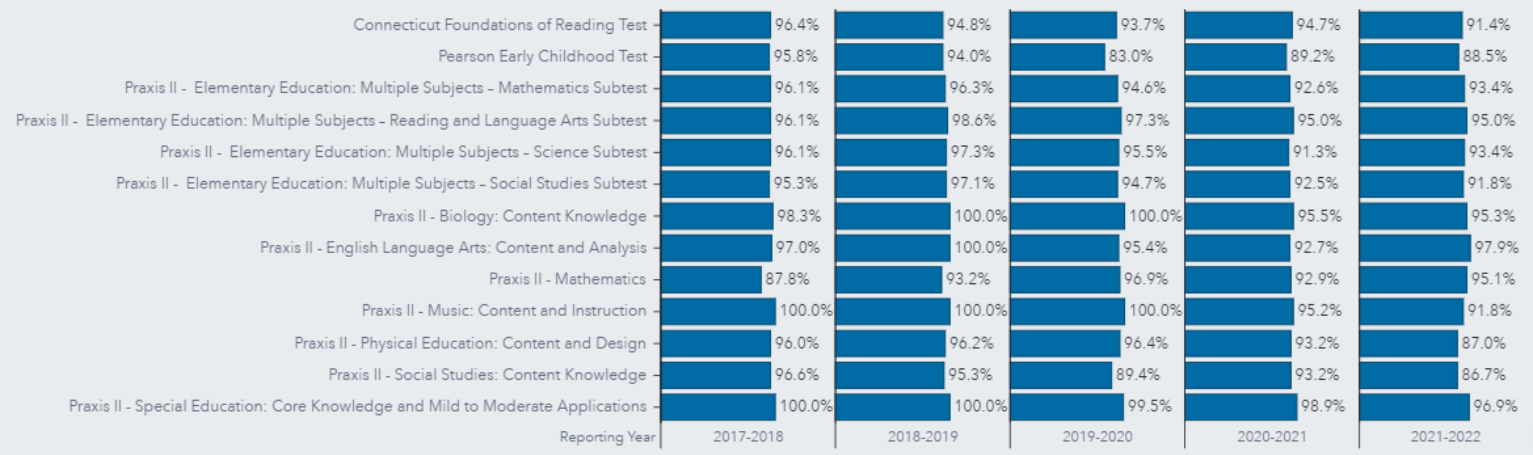
**2020-21 data are provided because the institution reported fewer than 5 completers in 2021-22.*

It is extremely important to note that a candidate may make their first attempt at a licensure test well in advance of completing the EPP, and sometimes even prior to *enrolling* in an EPP. For instance, of the 3,312 tests taken by the 2021-22 completers, 1,371 tests (41.4%) were taken by June 30, 2021, prior to the start of the final 2021-22 academic year. Therefore, the Best Pass Rate is a more accurate reflection of student performance on the licensure assessments after program completion.

The CSDE took a closer look at the individual content area assessments. Over 50% of the assessments taken by 2021-22 completers are in the four elementary education subjects of Reading and Language Arts, Mathematics, Science, and Social Studies. Approximately 20% are in the Connecticut Foundations of Reading Test.

The CSDE then explored the variability in pass rates among different content areas assessments. Any assessments with at least 40 tests attempted were included (Figure 1). Pass rates in 2021-22 exceeded 90% for almost all assessments with the exception of the Pearson Early Childhood Test, Praxis II – Physical Education: Content and Design, and Praxis II – Social Studies: Content Knowledge.

Five-Year Trend in Best Pass Rates for Licensure Assessments



With the awareness of these descriptive trends, the CSDE engaged Boston University's Wheelock Educational Policy Center (WEPC) to take a closer look at the effect of failing the first attempt of a required licensure exam on the likelihood of becoming a teacher in Connecticut. The policy brief produced by WEPC is attached in Appendix A. Key highlights from the WEPC study are presented below.

- Among individuals who "just pass" a Praxis II test on their first attempt, 82% go on to obtain a teaching certification, compared to 75% of those who fall just short of receiving a passing score.

- If passing cut-scores are reduced slightly (by approximately 4 points or 0.5 standard deviations), the number of Praxis II test takers obtaining certification would likely increase by about 25 individuals each year.

In June 2022, Governor Lamont allocated a total of $2 million dollars of federal, state-level reserve American Rescue Plan Act, Elementary and Secondary School Emergency Relief (ARP-ESSER) funding to provide financial support to teacher candidates. Grant funds were dispersed to each EPP in Connecticut and could be used to pay for candidate fees for licensing assessments including re-takes, certification fees, and costs associated with fingerprinting and background checks. As of February 20, 2024, approximately $1.35 million dollars were yet to be requested by the EPPs. Per federal rules, all remaining funds must be obligated (i.e., committed) by September 30, 2024. If all these funds are obligated by September 30, 2024, but not liquidated by the federal deadline of January 28, 2025, the CSDE may be able to request an extension from the U.S. Department of Education to continue

liquidating these funds until March 31, 2026. This would allow EPPs to continue utilizing these funds to offset the fees for licensure assessments that are incurred by aspiring educators in Connecticut's EPPs.

Discussion

The CSDE analyses illustrate that despite the decline in pass rates after the pandemic, more than 90% of all candidates and 85% of candidates of color continue to pass in one or more attempts. The WEPC analyses show that even among candidates who just fail a Praxis II on their first attempt, 75% go on to obtain a teaching certification. The Praxis II passing standards appear to rightly differentiate the "just-qualified" candidate from the "not-quite qualified" candidate, thus accurately evaluating **minimum content knowledge** in the subject area for individuals entering the teaching profession. Given the rigors of the CT Core Standards, it is prudent for Connecticut to maintain the expectation that new teachers will possess baseline content knowledge in the content area(s) they seek to teach.

However, Connecticut currently only offers one method – a timed, standardized test – to demonstrate this content knowledge. While the test is definitely valid for its intended use (i.e., to evaluate minimum content knowledge), and the vast majority of EPP completers (above 85%) are passing the assessments, it can be extremely discouraging for those candidates who have devoted multiple years in an EPP to learn that they are unable to demonstrate their content knowledge through the standardized test. Therefore, in lieu of requiring all candidates who are unable to pass a content area assessment after two attempts to continue retaking the exact same assessment, Connecticut should consider offering alternatives for those candidates who are close to passing Praxis II after two attempts, to demonstrate their minimum content knowledge.

At the CSDE's request, WEPC shared some approaches to multiple measures being considered by other states. Multiple states allow for alternatives if the candidate's score is within one standard error of measurement of the passing score (e.g., Alabama, Massachusetts, Missouri). Some consider more limited certification options if the candidate is missing a requirement (e.g., New Jersey) while others consider some combination of EPP recommendation, EPP grade point average, state-approved training, or approved coursework (e.g., Alabama, California, Massachusetts, Oregon).

The CSDE is particularly exploring two alternatives in some content areas for candidates who "just failed" the test after two attempts (i.e., scored within 1 standard error of measurement of the passing

score on either attempt). These are similar to the approaches currently being implemented in Massachusetts:

1. Re-take a lower cost "flex-test" option; or

2. Require the EPP to attest to the minimum content knowledge of the candidate in the relevant content area by providing evidence such as course grades or grade point average in specific courses.

In addition to offering alternatives for those who just-fail Praxis II after two attempts, the CSDE should also support existing EPPs to ensure that their course curriculum pathway will indeed prepare candidates with the adequate content knowledge to meet the "just-qualified" standard for their respective Praxis II assessment(s). The variability in pass rates among EPPs suggests that some remediation approaches may be more effective than others.

The CSDE will review Connecticut's cut scores for all Praxis II tests and ensure that they are not higher than the ETS recommended cut scores from standard-setting procedures.

The CSDE will also raise the findings from this study for discussion at the recently convened Connecticut Educator Certification Council. This Council includes broad representation from different groups including legislators, superintendents, teacher unions, EPPs, state and local boards of education, and other affected parties. The objective of this Council is to evaluate the effectiveness of current regulations with an eye to removing barriers for current and aspiring educators, and to modernize the educator workforce.

Conclusion

The results of this study indicate that while the vast majority of EPP completers are passing content area licensure assessments, some are not passing. While the Praxis II is a valid measure of minimum content knowledge for individuals entering the teaching profession and is warranted given the rigors of the CT Core Standards, alternative approaches to demonstrating minimum content knowledge should be considered for those who were close to passing Praxis II after two attempts. These alternatives can include a low-cost flex-test option, and attestation by the EPP. At the same time, the CSDE should also help EPPs review their course pathways to ensure that candidates are better prepared to demonstrate

that minimum content knowledge upon program completion. CSDE should also confirm that its Praxis II cut scores are not higher than the ETS recommended cut scores from standard-setting procedures.

Appendix A

Licensure Tests and Teacher Supply in Connecticut by Alexis Orellana and Marcus A. Winters. Boston University Wheelock College of Education & Human Development, Wheelock Educational Policy Center.

BU
Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center
WEPC

Licensure Tests and Teacher Supply in Connecticut

Alexis Orellana and Marcus A. Winters

EXECUTIVE SUMMARY

The Connecticut State Department of Education (CSDE) partnered with researchers from Boston University's Wheelock Educational Policy Center (WEPC) to examine the relationship between the state's teacher licensure policies and the composition and quality of its teacher workforce. This research project focused on understanding the association between performance on Praxis II subject-matter tests, which prospective teachers must pass in order to obtain licensure, and 1) their later effectiveness at improving student test scores, as well as 2) the overall supply of teachers into the workforce. This report summarizes key context, methods, results, and implications for consideration by policymakers. Overall, WEPC researchers found that Connecticut's current set of licensure test requirements are influential in shaping the composition of the teacher workforce. Separately but relatedly, WEPC found a weak link between licensure test performance and teacher impacts on student test score growth.

A Technical Appendix with more detailed information about this research study can be found at wheelockpolicycenter.org/all-research.

STATE AND NATIONAL RESEARCH CONTEXT

If you want to teach in a U.S. public school, chances are you will have to take and pass at least one standardized licensure exam. In fact, before the pandemic, all 50 states required prospective teachers to pass one or more subject-matter tests to obtain teaching certification, and 15 states required applicants to pass basic skills tests to gain admission into a teacher preparation program. While ubiquitous, these testing requirements are also controversial, and debates about their purpose and impact have sparked anew in recent years given growing concern about teacher shortages and the lack of diversity in the educator workforce.

Like many other states, Connecticut has been evaluating and revising its policies within this national context. In 2016, for example, Connecticut was one of the first states to do away with the requirement that prospective teachers take and pass a basic skills test (Praxis I) for entry into a teacher preparation program. Then, in 2021, amidst continued change efforts underway by CSDE, legislation in Connecticut (SB 1202) required the department of education to assess requirements governing content-area mastery and consider a multiple measures approach. Aligned with this charge, CSDE engaged with Boston University's WEPC to investigate the relationships between these licensure tests, student achievement, and, to the extent possible, the composition and quality of the teacher workforce.

This research effort builds on and extends studies conducted in several other states by examining the relationships between licensure test performance and student achievement in Connecticut specifically. In states like North Carolina, Washington, and Massachusetts, previous work by other researchers revealed present, though weak, correlations between test scores and later teacher impacts.[1] Replicating these analyses was an important first step for WEPC's research in Connecticut, as it provided insight into the predictive utility of the tests themselves within the state's own context. The next line of inquiry, however, was more novel within the broader research base, shedding new light on the ways in which licensure test requirements affect the entry of prospective educators into the teaching profession.

Licensure tests are not unique to the teaching workforce. About 30% of U.S. workers are employed in an occupation that requires a government license.[2] Within some of these other professions, there is growing evidence to suggest that increases in licensure requirements may restrict the supply of workers (e.g., physical therapy, accountants, cosmetology).[3] This study adds to this research base by providing further information about the effects of occupational licensure requirements on workforce entry within the specific context of education.

DATA AND APPROACH

In this study, the authors examine longitudinal data from Connecticut, including all Praxis II scores beginning in 1995, teacher certification data from 2002 to 2022, staff assignment data from 2002 to 2020, and student achievement data from 2014 to 2021. The study sample consists of about 85,000 individuals with Praxis II scores; about 51,000 of them also have teacher employment records within the state.

There are a few important things to note about the data and analytic approach used in this study:

- Praxis II is **not a single test** but rather a wide range of subject-specific assessments. Each unique subject-area license requires different tests, some requiring one, and others, like elementary, requiring multiple. WEPC's analysis accounts for differences across tests, but many tests have too few participants for the authors to be able to confidently identify their distinct impacts on workforce entry separate from the overall pattern observed for all Praxis II tests.[4] In the case of Praxis II tests in STEM subjects, the authors were able to detect a distinct effect of first-time failure on workforce entry; however, this STEM-specific finding was similar in magnitude to the overall finding for all subjects.
- The researchers observe all test-taking attempts (first attempt, best attempt, and any retakes) but **primarily report the results for each test-taker's first attempt**, as they include the largest sample of test-takers and results follow a largely consistent pattern across attempts.
- When examining the relationship between Praxis II subject-matter test performance and teachers' later effectiveness at improving student test scores, the authors **look only at the tests that are directly tied to the grades and subjects in which students take state standardized assessments** (i.e., third-through-eighth grade math and ELA). This means that they do not estimate the predictiveness of Praxis II performance for, e.g., social studies, science, music, art, or high-school math and English teachers, as these roles are not tied to state-tested grades and subjects. Additionally, this analysis does not examine whether licensure test performance predicts teacher effectiveness at improving non-tested outcomes, which are undoubtedly important though less directly tied to the content matter found in Praxis II tests.
- In attempting to establish the relationship between Praxis II performance and later teacher effectiveness, the authors include only individuals that end up teaching in a Connecticut public school in a tested grade and subject area. Thus, the results may not directly apply to those who took a Praxis II test but were never observed as a teacher in a state-tested grade and subject. However, in examining the impact of the licensure tests on the overall supply of the workforce, the authors include all individuals taking any Praxis II test, whether they show up in the workforce or not. Because of this difference in which test-takers are included in each analysis, the researchers **caution against too many generalizations between the two different sets of results identified in this study**.
- Because of data availability limitations at the time this report was written, the researchers were **unable to examine variations based on prospective teachers' race/ethnicity or gender**. Future work may be able to incorporate additional data and/or analytic methods to address the critically important question of whether the effects of licensure test requirements on workforce entry differ for teachers from different backgrounds.
- This study focuses on Praxis II assessments and **does not include the Foundations of Reading test** that is also administered as part of Connecticut's teacher licensure process.

Estimating the relationship between licensure scores and later teacher effects on student achievement. Similar to prior studies, the researchers apply a value-added model to estimate for each teacher the difference in the average test scores of students they instruct and the score that these students would be predicted to achieve based on their prior year test scores and other observed

characteristics. They then examine the correlation between these value-added scores in either math or ELA and a teacher's relevant licensure test scores.

Estimating the causal effect of failing the licensure test on becoming a teacher. An important feature of this study is the ability to measure the causal link between a prospective teacher's success or failure on a licensure test and their eventual entry into the profession. The researchers used a regression discontinuity design to exploit the similar attributes of candidates right above and below the passing cut-score for each test and examine the differences in their trajectories into or away from the profession.

Full details on the study design, sample, and methodology are available in the Technical Appendix at wheelockpolicycenter.org/all-research.

RESULTS

Overall, the research team found that:

- Consistent with other studies from other states, there is a positive though very small relationship between teacher candidates' performance on Praxis II subject-matter tests and their later impacts on student achievement, as measured by test score value-added in ELA and math.
- In Connecticut, failing a first attempt at a Praxis II subject-matter test significantly and substantially reduces a candidate's likelihood of moving forward to become a certified public-school teacher.[5]
  - Among individuals who just pass a Praxis II test, 82% go on to obtain a teaching certification, compared to 75% of those who just fall short of receiving a passing score.
  - In a scenario where passing cut-scores are reduced slightly (by approximately 4 points or 0.5 standard deviations), the number of Praxis II takers obtaining certification would likely increase by about 25 individuals each year. This estimate is based on various assumptions, which may not hold true in reality, but it provides a sense of the scale of test-takers affected by the Praxis II requirement.

FUTURE WORK

These results do not suggest an obvious path forward. There is a lot we are continuing to learn about the forces that move the teacher labor market, and more we must understand about the role licensure tests play in a complex policy environment. In most cases, states use licensure tests as a way to ensure that teachers entering the profession possess the minimum content knowledge needed to teach specific grades and subjects. The question policymakers are continually seeking to understand, however, is whether these tests are serving this intended purpose and whether there are unintended consequences associated with them.

The findings in this report shed some light on the potential trade-offs at play in requiring prospective teachers to pass licensure tests. On the one hand, the study finds that Connecticut's Praxis II requirement is deterring some prospective teachers who fail their first attempt at a subject-matter licensure test from persisting into the profession. On the other hand, there is ambiguity associated with predicting what would happen to the overall supply and entry of teachers into the workforce if these requirements were changed or removed. It is also important to consider that, in at least some grades and subjects, these tests are not highly predictive of teachers' later impacts on student test score growth. State and local leaders in Connecticut and other places will need to make judgment calls about whether and how to adapt existing licensure requirements to account for this evidence alongside their broader goals for increasing teacher supply, diversity, and quality. As policymakers consider these potential policy changes, there are several points worth bearing in mind:

1. **We cannot fully predict the impact of changing licensure or testing requirements on the composition of the teacher workforce.** Based on the results of this study, we cannot say with certainty what would happen if testing requirements were significantly reduced or altered. There are still too many unknowns about the various conditions that influence and intersect in

an individual's decisions to enter and persist in the teacher pipeline. For example, our analysis does not consider the extent to which having a licensure requirement alters the pool of individuals who pursue a career in teaching. That said, in the context of the pandemic, many states have significantly altered their licensure requirements, particularly around testing, creating natural experiments in some areas that researchers are seeking to document and understand. WEPC is one such entity, partnering with Massachusetts, for instance, to understand the impacts of their pandemic-induced emergency license. So, while we don't yet have solid evidence on the counterfactuals of licensure exams as they have traditionally been used, we should have new insights to lend to this understanding within the coming months and years.

2. **Where you set the passing cut-score likely matters.** There is immense discretion in where states decide to set the pass/fail bar for these tests. In the case of the Praxis suite, most states have adopted the ETS-recommended cut score, and in so doing have also helped to maintain consistency across state borders for the purposes of licensure reciprocity. The results of this study, however, suggest that even small changes in where states set the passing score are likely to affect the number of individuals entering the teacher workforce. Thus, states should closely evaluate whether their chosen licensure test cut-scores truly represent their minimum content knowledge expectations for novice teachers.

3. **It is worth continuing to question whether we have the right measures in place to assess teacher and student knowledge.** The premise behind licensure test requirements is that teachers' content knowledge matters for their ability to effectively help students learn that content, and that licensure tests are an accurate predictive tool for measuring this content knowledge. If we believe these two things to be true, we would expect to see a relatively strong relationship between licensure test performance and student test score growth in aligned subjects. Since in this case we don't observe this strong relationship, it may mean we should be interrogating the underlying assumptions behind what licensure tests measure and how. To this point, several states have begun exploring other ways to assess teachers' baseline content knowledge—for example requiring a degree in the subject area or a portfolio review—but not enough evidence exists to date to suggest that these alternatives are any better aligned with later impacts on student learning. Unfortunately, until better options become available, states are in the unenviable position of weighing the tradeoffs between the current system, a lack of other viable alternatives, and the unknowns about what happens without a clear requirement in place.

ENDNOTES

1. Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources, 41*(4), 778-820; Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682; Clotfelter, C. T., Ladd, H. F., & 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; Goldhaber, D., & Hansen, M. (2010). Race, gender, and teacher testing: How informative a tool is teacher licensure testing? *American Educational Research Journal, 47*(1), 218-251.
2. Kleiner, M. M., & Krueger, A. B. (2013). Analyzing the extent and influence of occupational licensing on the labor market. *Journal of Labor Economics, 31*(S1), S173-S202.
3. Adams, A. F., Jackson, J. D., & Ekelund, R. B. (2002). Occupational licensing in a "competitive" labor market: The case of cosmetology. *Journal of Labor Research, 23*(2), 261-278; Cai, J., & Kleiner, M. M. (2020). The labor market consequences of regulating similar occupations: The licensing of occupational and physical therapists. *Journal of Labor Research, 41*, 352-381; Jacob, J., & Murray, D. (2006). Supply-side effects of the 150-hour educational requirement for CPA licensure. *Journal of Regulatory Economics, 30*, 159-178.
4. Specifically, the following list of Praxis II tests was considered: English Language, Literature, and Composition: Content Knowledge (41); English Language, Literature, and Composition: Essays (42); English Language, Literature, and Composition: Content and Analysis (44, 5044); Middle School English and Language Arts (49, 5047, 5049); Elementary Education: Multiple Subjects–Reading and Language Arts Subtest (5002, 5032); English Language Arts: Content and Analysis (5039); Mathematics (60); Mathematics: Content Knowledge (61, 5061, 5161); Middle School Mathematics (69, 5169); Elementary Education: Multiple Subjects–Mathematics Subtest (5003, 5033).
5. Teacher certification is defined as being associated to a teaching endorsement (i.e., discarding endorsements such as driver education, licensed practical nurse, librarian, speech and language pathologist, school psychologist, etc.) plus having any of the following certifications: "Initial Educator," "Permanent Teaching", "Professional Educator," or "Standard Teaching."

ACKNOWLEDGEMENT

WEPC is grateful to the Connecticut State Department of Education for providing data access, feedback, and comments throughout the development of this report, with particular thanks to Shuana Tucker, Ajit Gopalakrishnan, and David Alexandro.

OUR MISSION

The Wheelock Educational Policy Center (WEPC) conducts and disseminates rigorous, policy-relevant education research in partnership with local, state, and federal policymakers and stakeholders to improve educational opportunities and holistic outcomes for underserved students.

www.wheelockpolicycenter.org

wheelockpolicy@bu.edu

Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

Licensure Tests and Teacher Supply in Connecticut:

Technical Appendix

Alexis Orellana and Marcus A. Winters

Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

1 Data

1.1 Licensure Tests

Similar to other states, in Connecticut the typical certification process requires an applicant to complete a state-approved educator preparation program and pass the subject-specific tests required to obtain an endorsement in their area of specialization. During our sample period, the state employed tests related to both of these certification requirements, all of which were created and administered by Educational Testing Service (ETS). Minimum passing scores for each test are determined by the Connecticut State Department of Education (CSDE).

We observe records for all licensure tests submitted to CSDE each year from 1995 to 2021. ETS routinely submits to CSDE all scores from test-takers who list Connecticut as their state of residence, take the test in Connecticut, or specify a preference for their scores to be submitted there. Each record contains an individual identifier, test-type identifier, score, and date. This information allows us to observe and distinguish each administration and test taken by each candidate during the sample period. Unfortunately, we do not observe demographic characteristics, such as gender or race, for all test-takers because ETS does not report such information to CSDE as part of the score transfer.

During the hiring process, schools observe a candidate's certification and endorsement status, and thus can infer that a candidate has passed the necessary licensure tests. However, schools do not typically observe an applicant's specific licensure test score(s) or information about the number of attempts the candidate required to pass.

1.1.1 Subject-Matter Certification Test: Praxis II

We focus our analysis on the various forms of Praxis II, also known as Praxis Subject, which assesses knowledge of specific subjects, as well as general and subject-specific teaching skills. Candidates typically take these tests during the final year of their preparation program as part of applying to obtain a teaching certification or endorsement to teach a particular subject.

Each of the several subject-matter tests is linked to a particular endorsement. Table 1 shows the link between some of the endorsement codes offered in Connecticut and the Praxis II tests required. Some endorsement codes involve passing more than one test (for example, *Elementary Grades, K-6*). In these cases, we group all sub-tests and employ the minimum score as the forcing variable in the analysis described in Section 4.[1]

Table (1) Praxis II Tests and Teaching Endorsements in Connecticut

Endorsement	Description	Praxis II Test	Additional Test
13	Elementary Grades K-6	5002 + 5003 + 5004 + 5005	Foundations of Reading
15	English 7-12	44, 49 or 5039	
26	History/Social Studies 7-12	81 or 5081	
29	Mathematics 7-12	61 or 5161	
30	Biology 7-12	235 or 5235	
31	Chemistry 7-12	242 + 245 or 5245	
32	Physics 7-12	262 + 265 or 5265	
33	Earth Science 7-12	571 or 5571	
34	General Science 7-12	433 + 435 or 5435	
47	Technology Education PK-12	51 or 5051	
49	Music PK-12	111+ 113 or 114 or 5114	
111	TESOL PK-12	361 or 5362	
165	Comprehensive Special Education K-12	543 or 5543	Foundations of Reading
215	English Middle School 4-8	5047	
226	History/Social Studies Middle School 4-8	89 or 5089	
229	Mathematics Middle School 4-8	69 or 5169	
230, 231, 232, 233, 234, 235	Middle Grades Science	5540	
305	Elementary Grades 1-6	5032 + 5033 + 5034 + 5035	Foundations of Reading

Notes: This table presents the Praxis II test requirements to earn a teaching certification in Connecticut. We employ this correspondence to identify whether applicants obtained a certification in the same Praxis II subject. The first and second columns display the code and subject-area description of each endorsement. The third column details which Praxis II tests are required in each case. The last column indicates whether an additional test (Foundations of Reading) is also required. This additional test is not used in our analyses since it is not administered by ETS.

1.2 Certification Data

We link applicants' scores on licensure tests to Connecticut's certification data between 2002 and 2021. For each person who applied to the state for certification and/or endorsement these records contain the certificate type, the date when the certification was issued, and the endorsement code indicating the subject in which the license grants the teacher permission to instruct. In addition, these data also include basic demographic information for those applying for certification, including the candidate's race/ethnicity and gender.

For our analyses, we define a certified teacher as one who has obtained a renewable Initial or Provisional Educator Certificate.[2] In order to gain an Initial Educator Certification in the state, in addition to passing the relevant Praxis II test, an individual must hold a bachelor's degree, complete required coursework in professional education, general education, in some cases complete a subject-area major, and provide a recommendation for certification from a state-approved program. Once they believe they have fulfilled the requirements, individuals apply for certification by creating an account on the Connecticut Educator Certification System and paying a nominal fee. Since obtaining a certification requires an individual to actively apply and demonstrate that they have completed necessary benchmarks implies that those who hold a certification have some interest in obtaining a teaching position beyond what is evidenced by simply passing the licensure test, we consider it to be a reasonable proxy for seeking a teaching position.

We separately distinguish those who teach on a nonrenewable Interim Educator Certificate or permit to teach within a shortage area.[3] Though all teaching within a Connecticut public school should have one of these certification types, we observe a small number of teachers with valid initial licensure scores who we do not match to a license.

1.3 Employment Records

We observe staff assignment data in all Connecticut public schools between 2002 and 2020. These records contain a unique Educator Identification Number (EIN), school code, position, and, in the

case of teachers, the subject taught. We use the EIN identifiers to match teachers' information across datasets. Additionally, we employ these records to estimate the effect on the likelihood of observing an applicant serving as a teacher for at least five years.

1.4 Additional Teacher and Student Administrative Data

Our analysis describing the relationship between scores on licensure tests and a teacher's later impacts on students requires data matching students to teachers within the state over time. Student-level data contains test scores, demographic characteristics, and participation in programs such as special education and English language supplemental services. We use course offerings and student-course-grade information to construct a classroom identifier and link students to their teachers.

When estimating teacher value-added we restrict the analysis to the set of classrooms assigned to educators with a valid identifier. In addition, we only consider classrooms linked to one teacher during the corresponding school year. This restriction is necessary to correctly identify each teacher's contribution in our analysis.

We link teachers to students with valid test scores in Language or Math in grades 3 through 8 for each year from 2014-15 through 2020-21, except for 2019-20, when students did not take the test due to the Covid-19 pandemic. We successfully matched 95% of students to a single classroom teacher.

2 Summary of Data Matching Process

To link students and teachers we match the TCS course offering and TCS student course grade datasets. To merge both datasets, we construct a classroom identifier as a unique *year-school-NCES code-section-start date-end date* combination. We do not consider year 2019-20 when the Smarter Balanced assessments were not administered.

We restrict our analysis to the set of classrooms assigned to educators who hold a valid

ein identifier. In addition, we only consider classrooms linked to one teacher during the corresponding school year. This restriction is necessary to correctly identify each teacher's contribution in our analysis.

Table 2 summarizes the number of students taking the math and language Smarter Balanced assessments (SB) between 3rd and 8th grades and the proportion of students we successfully match to a classroom with one teacher.

Table (2) Distribution of matches between 2014-2020

Year	Not Matched	Matched	Students taking ELA or math SB tests
2014	5.2%	94.8%	235,497
2015	3.1%	96.9%	234,993
2016	2.8%	97.2%	234,759
2017	3.1%	96.9%	233,465
2018	5.7%	94.3%	231,109
2019			0
2020	13.9%	86.1%	214,291
Total	5.5%	94.5%	1,384,114

3 Estimating Relationship Between Licensure Scores and Value-Added

We use a two-step approach to investigate the relationship between a teacher's score on the respective Praxis exam and their later impact on student achievement. The first stage estimates the teacher's independent contribution to student test scores, on average, commonly referred to as the teacher's "value-added". The second stage then measures the association between the teacher's value-added as estimated from the first stage and their Praxis score.

For the first-stage analysis, we use a conventional value-added approach to produce an

estimate for each teacher's impact on student test scores. The general model takes the form:

$$y_{ijst} = X'_{ijst}\beta + f(y_{ijst-1})\lambda + \phi_j + \epsilon_{ijst} \quad (1)$$

Where y_{ijst} is the test score for student i instructed by teacher j within school s during year t; X is a vector of student and classroom characteristics and grade fixed effects; $f(y_{ijst-1})$ is a cubic function of the student's test score at the end of the previous year in math and language; ϕ_j is a teacher fixed effect; ϵ_{ijst} is a stochastic term; and β and λ are parameters to be estimated.

For each teacher we capture $\hat{\phi_j}$, which is our estimate for each teacher's contribution to student test scores conditional on the other covariates. A common challenge with value-added approaches is that for any individual teacher, the sample size used to identify the relevant fixed effect may be quite small and hence estimated with a substantial degree of noise. We address this issue by employing the Bayesian Shrinkage adjustment, as is typical in the value-added literature.

That we control for prior test scores focuses the model on estimating teacher impacts on student test scores gains. We employ a cubic function for lagged test scores in order to allow for differences in expected growth for students at different points on the distribution of prior test scores. Prior research demonstrates that value-added models that account for prior test scores appear to be forecast unbiased when applied within large-scale administrative data.

Equation 1 represents our base value-added model, which we use as our primary estimate for teacher value-added for which we report results in the main body of the paper. However, results are similar from models that incorporate various fixed effects for schools or school-by-year and from models that remove the function for prior test scores and rather incorporate a student fixed-effect.

For the second step in the analysis, we aggregate the data to the teacher level and estimate a regression where the dependent variable is the teacher's estimated value-added from the first stage, $\hat{\phi_j}$, and independent variables include the teacher's score on the respective Praxis exam (P_j) and a vector of time-invariant teacher characteristics (gender, race/ethnicity, and education level

indicators) represented by Z'_j. Formally:

$$\hat{\phi}_j = P_j\gamma + Z'_j\delta + \eta_j \tag{2}$$

We are primarily interested in the estimate for γ, which represents the relationship between the teacher's score on the Praxis II exam and their estimated value-added contribution to student test scores. We estimate equation 2 separately for Language, Math, and elementary teachers.

Figure 1 shows the association between licensure score and our empirical Bayes estimates of test score value-added by subject and test-type. For both subjects, we find a small, positive relationship. For ELA teachers, a standard deviation increase in Praxis II score associates with a gain of 0.0052σ in test scores. For math, a standard deviation increase in Praxis II scores associates with a gain of 0.0081σ.

4 Estimating the Causal Effect of Failing a Licensure Test on Progressing Toward Becoming a Teacher

Our goal is to estimate the causal effect of an individual failing their first attempt on a licensure test on their pathway to becoming a teacher. A naive comparison is likely biased by unobserved differences related to the likelihood of failing and one's trajectory towards becoming a public school teacher. We overcome this challenge by leveraging the sharp discontinuity in passing a given test that occurs at the designated cutoff.

Let i denote an applicant taking test j for the first time. Each test j has a minimum passing score $\bar{x}_j$. We center scores around the corresponding cutoff and standardize them using the within-sample standard deviation.[4] We denote this variable x_{ij}. When a test j considers more than one subtest, we define x_{ij} as the minimum value across all sub-tests. We account for changes in the tests over time and differences between different subject-area tests by including fixed ef-

fects for year and specific test administered. Our main analyses are based on a sharp regression discontinuity design using the following specification:

$$y_{ij} = \alpha + f(x_{ij}) + \beta \mathbb{1}(x_{ij} < 0) + \phi_j + \phi_t + \epsilon_{ij} \tag{3}$$

The term $f(x_{ij})$ is a parametric function of the (normalized) score obtained by applicant i, which our primary model employs as quadratic and allows for changes in the slope at the cutoff value.[5] We estimate local linear regressions to observations that fall within optimal bandwidths from the cutoff as calculated using the methodology of Calonico et al. (2014) (hereafter, CCT). Our primary results are from models that employ a triangular kernel. The sample includes an individual's first observed score on the relevant licensure test, excluding first-time test-takers who we previously observe teaching within a Connecticut public school. This latter exclusion should account for current teachers whose first attempt took place in a year prior to our data beginning.

The key identifying assumption for β is that the relationship between a candidate's score and the outcome would be smooth at the passing threshold if not for the fact that scoring above the line satisfied the passing requirement. There are two particular threats to this assumption. The first is the potential for individuals to manipulate their scores around the cutoff. The institutional features of the certification process in Connecticut make violating this assumptions unlikely. Figure 2 shows no indication of manipulation around the cutoff values. We find no statistical evidence to reject the null hypothesis of continuity around the passing threshold. The p-value of the discontinuity test is 0.35.

The second threat to identification is the potential for discontinuities in the value for confounders around the threshold. To investigate the potential for this threat, authors typically look for balance in the value of observed baseline characteristics on either side of the threshold. Unfortunately, such conventional balance tests are not available to us because we observe demographic information only for individuals who apply for a certification or endorsement. Nonetheless, given the nature of the tests we argue that it is highly unlikely for there to exist a systematic discontinuity

at the passing threshold in the characteristics of test-takers.[6]

Table 3 reports regression discontinuity estimates for the effect of failing the first administration of Praxis II. For those scoring at the threshold, failing the first administration of Praxis II reduced the likelihood that a candidate obtained any teaching certification by about 6.6 percentage points. Failing the test required for endorsement to teach a STEM subject reduced the likelihood of obtaining the endorsement by about 8.9 percentage points. And failing the test for endorsement to teach within special education reduced the likelihood of obtaining that endorsement by 10.7 percentage points.

Table (3) RD Estimates for Effect of Failing First Administration of Licensure Test

	(1) Any Certification	(2) STEM	(3) Special Education
Failed Praxis II	-0.066*** (0.013)	-0.089*** (0.033)	-0.107** (0.042)
Average Outcome	0.79	0.68	0.82
Bandwidth	(-0.58,0.78)	(-0.51,0.69)	(-0.63,0.69)
N	34,307	6,207	3,425

Notes: This table presents estimates of the effects of failing the first attempt at Praxis II on the likelihood of eventually obtaining any teaching certification, obtaining an endorsement to teach within a STEM subject, and endorsement to teach special education. Analyses of STEM and special education endorsements are restricted to the first administration of a test associated with that particular endorsement, rather than the first Praxis II attempt. Bandwidths are selected following Calonico et al. (2014) and reported at the bottom of the respective analysis. Each regression controls for the difference between the individual's licensure score and the passing score for the respective test within a linear function allowing for changes in the slope at the threshold, as well as both year and test fixed effects. Heteroskedastic robust standard errors reported in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure (1) Association between Praxis II scores and Teacher Value-Added

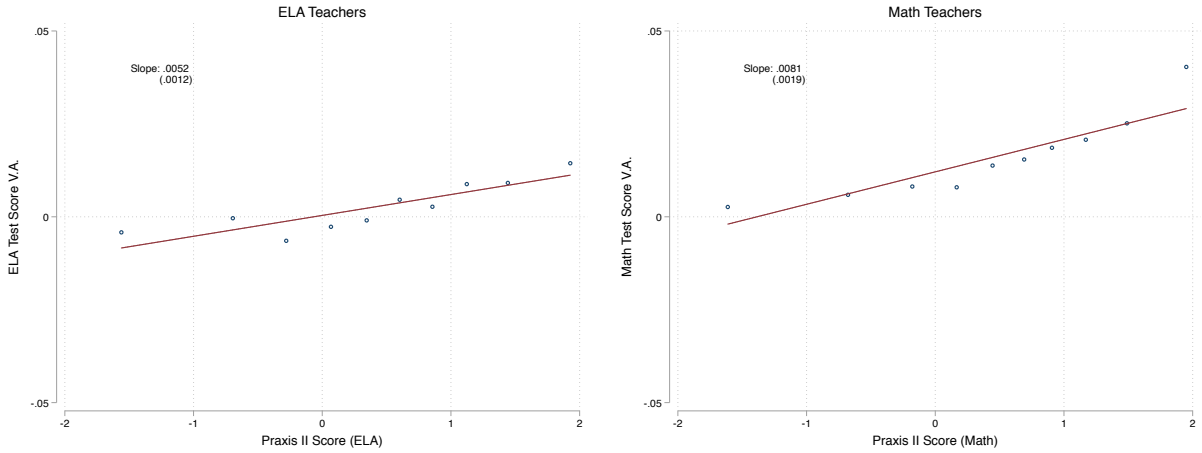
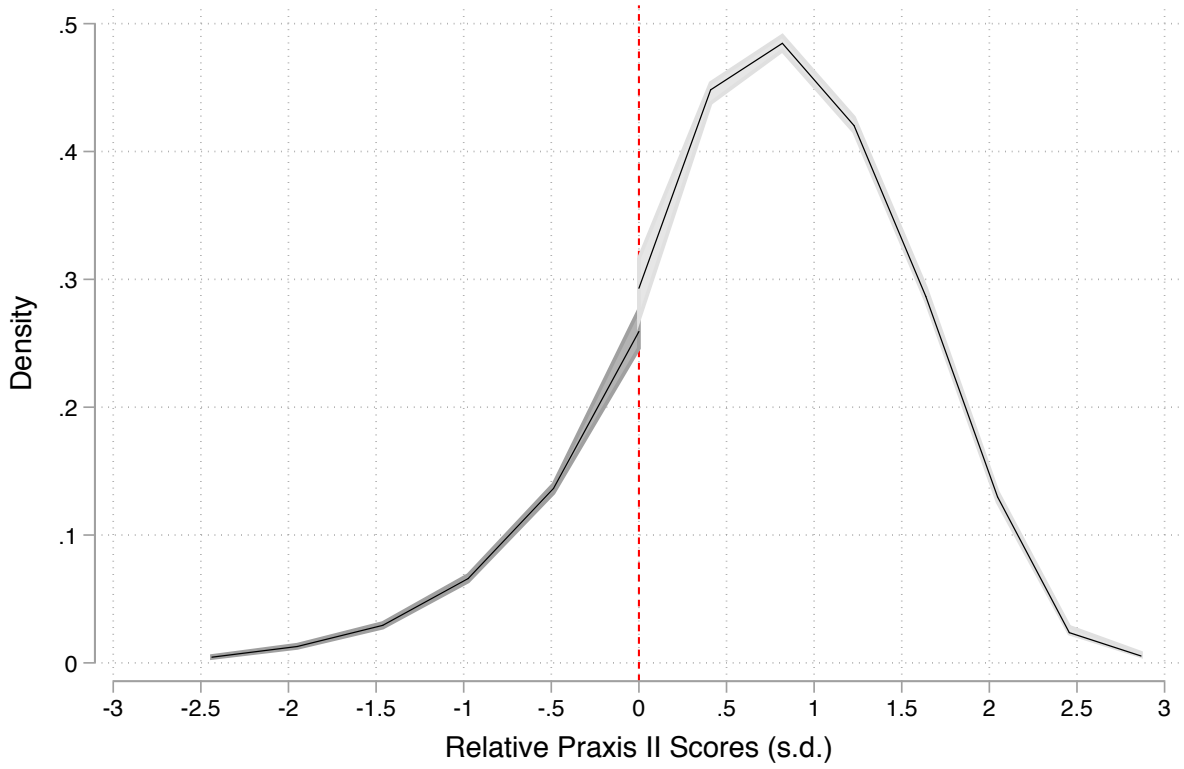


Figure (2) Density of Praxis II tests around the threshold



Notes: This figure illustrates the density of standardized Praxis II scores around the threshold. The density and 95% confidence intervals at each side of the cutoff were estimated following Cattaneo et al. (2018). The discontinuity test has a p-value of 0.35. These values imply there is no statistical evidence to reject the null hypothesis of no discontinuity at the threshold.

Notes

[1]Table 1 shows a few endorsements require an additional test, *Foundations of Reading*, which is not administered by ETS. We do not consider this subtest in our analyses.

[2]An Initial Educator Certificate is a 3-year certificate for those who have either completed a preparation program or have at least 20 school-months of teaching experience in a non-public school. A Provisional Educator Certificate is an 8-year certification for who who have at least 10 school-months of experience under a different certificate type or at least 30 school-months of appropriate experience in a non-public school.

[3]An Interim Educator Certificate is a nonrenewable certificate issued to those who have not fully completed either the testing or coursework requirements to obtain an Initial Educator Certification.

[4]We employ standardized scores instead of raw scores because sometimes tests differ in their scale. For example, each applicant must approve two exams to earn an endorsement in Chemistry. The first one, *Chemistry: Content Knowledge*, is scored using 1-point intervals while *Chemistry: Content Essays* uses 5-point intervals.

[5]We estimate the model using the rdrobust command in STATA and report as our primary results estimates from the Robust specification.

[6]Goldhaber and Hansen (2010) employ balance tests to assess differences in race and gender between applicants who fail and pass Praxis II tests in North Carolina. They do not find evidence of discontinuities at any of the cut scores they analyze.

References

Calonico, S., Cattaneo, M., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82:2295–2326.

Cattaneo, M., Jansson, M., and Ma, X. (2018). Manipulation Testing Based on Density Discontinuity. *Stata Journal*, 18(1):234–261.

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.

OUR MISSION

The Wheelock Educational Policy Center (WEPC) conducts and disseminates rigorous, policy-relevant education research in partnership with local, state, and federal policymakers and stakeholders to improve pk–20 educational opportunities and holistic outcomes for underserved students.

www.wheelockpolicycenter.org
wheelockpolicy@bu.edu

Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

WEPC
WHEELOCK EDUCATIONAL POLICY CENTER