# Alternative Assessment Pilot Evaluation

Final Report

James Cowan, Dan Goldhaber, Roddy Theobald

Center for the Analysis of Longitudinal Data in Education Research (CALDER), American Institutes for Research

Dana Ansel

February 2025

Contents

[Executive Summary 1](#_Toc188372842)

[Data Sources 1](#_Toc188372843)

[Summary of Alternatives 2](#_Toc188372844)

[Key Findings 3](#_Toc188372845)

[Introduction and Background 6](#_Toc188372846)

[Candidate Experiences 9](#_Toc188372847)

[Communications and Literacy Skills (CLST) Alternatives 9](#_Toc188372848)

[MTEL-Flex 12](#_Toc188372849)

[Program Attestation 15](#_Toc188372850)

[Candidate Background and Identity 18](#_Toc188372851)

[Communications and Literacy Skills Alternatives 18](#_Toc188372852)

[MTEL-Flex 20](#_Toc188372853)

[Program Attestation 21](#_Toc188372854)

[Access to the Profession 23](#_Toc188372855)

[Licensure and Employment Data 23](#_Toc188372856)

[Empirical Methods 25](#_Toc188372857)

[Communications and Literacy Skills Alternatives 26](#_Toc188372858)

[MTEL-Flex 29](#_Toc188372859)

[Program Attestation 34](#_Toc188372860)

[Teacher Effectiveness 37](#_Toc188372861)

[Data and Empirical Methods 37](#_Toc188372862)

[Communications and Literacy Skills Alternatives 44](#_Toc188372863)

[MTEL-Flex 45](#_Toc188372864)

[Program Attestation 47](#_Toc188372865)

[Summary 49](#_Toc188372866)

[Considerations 50](#_Toc188372867)

[Appendix A. Program Attestation Interviews 1](#_Toc188372868)

[Appendix B. Additional Results 1](#_Toc188372869)

[References 6](#_Toc188372870)

Exhibits

[Exhibit 1. MTEL Alternatives Participation 4](#_Toc188372871)

[Exhibit 2. Cumulative CLST Alternative Submissions by Date 10](#_Toc188372872)

[Exhibit 3. CLST Alternative Submissions by Test Series and Subtest 10](#_Toc188372873)

[Exhibit 4. Cost and Format of CLST MTEL and Communications and Literacy Skills Alternatives 11](#_Toc188372874)

[Exhibit 5. Importance of Factors Influencing Decision to Take CLST Alternatives 12](#_Toc188372875)

[Exhibit 6. Cumulative MTEL-Flex Submissions by Date 13](#_Toc188372876)

[Exhibit 7. Candidate MTEL-Flex Retake Status by Month of Qualification 14](#_Toc188372877)

[Exhibit 8. MTEL-Flex Submissions by Subject Test 15](#_Toc188372878)

[Exhibit 9. Cumulative Program Attestations by Date 15](#_Toc188372879)

[Exhibit 10. Teacher Candidate Characteristics and CLST Alternatives Choice 19](#_Toc188372880)

[Exhibit 11. Teacher Candidate Characteristics and CLST Alternatives Choice 19](#_Toc188372881)

[Exhibit 12. Teacher Candidate Characteristics and MTEL-Flex Choice 20](#_Toc188372882)

[Exhibit 13. Teacher Candidate Characteristics and MTEL-Flex Choice 21](#_Toc188372883)

[Exhibit 14. Program Attestation Candidate Background 22](#_Toc188372884)

[Exhibit 15. Program Attestation Candidate Ethnoracial Identity 22](#_Toc188372885)

[Exhibit 16. Licensure and Employment Outcomes 24](#_Toc188372886)

[Exhibit 17. Timeline of Milestones Relative to First Communication and Literacy Skills Test Date 24](#_Toc188372887)

[Exhibit 18. Timeline of Milestones Relative to First Subject Test Date 25](#_Toc188372888)

[Exhibit 19. Candidate Outcomes on the Traditional CLST MTEL and CLST Alternatives 26](#_Toc188372889)

[Exhibit 20. Differences in Workforce Outcomes Adjusted for Candidate Background 28](#_Toc188372890)

[Exhibit 21. Candidate Outcomes on the MTEL-Flex 29](#_Toc188372891)

[Exhibit 22. Differences in Workforce Outcomes Adjusted for Candidate Background 30](#_Toc188372892)

[Exhibit 23. MTEL-Flex Pass Rates by Test Month 31](#_Toc188372893)

[Exhibit 24. Candidate Outcomes on the MTEL-Flex 32](#_Toc188372894)

[Exhibit 25. MTEL-Flex Qualification and Retesting 33](#_Toc188372895)

[Exhibit 26. Dynamic Effects of MTEL-Flex Qualification on Retesting 34](#_Toc188372896)

[Exhibit 27. Percentage of Program Attestation Candidates Attaining Licensure/Employment Milestones 35](#_Toc188372897)

[Exhibit 28. Differences in Workforce Outcomes Adjusted for Candidate Background 36](#_Toc188372898)

[Exhibit 29. Licensure and Employment Outcomes 39](#_Toc188372899)

[Exhibit 30. Effectiveness of Candidates Passing CLST Alternatives 44](#_Toc188372900)

[Exhibit 31. Effectiveness of Candidates Passing MTEL-Flex 45](#_Toc188372901)

[Exhibit 32. Effectiveness of Candidates Passing Program Attestation 47](#_Toc188372902)

[**Exhibit B–1.** Teacher Candidate Pass Rates on CLST Alternatives by Ethnoracial Identity 1](#_Toc188372903)

[**Exhibit B–2.** Teacher Candidate Pass Rates on MTEL-Flex by Ethnoracial Identity 2](#_Toc188372904)

[**Exhibit B–3.** Teacher Candidate Pass Rates on MTEL-Flex by Ethnoracial Identity 3](#_Toc188372905)

[**Exhibit B–4.** Teacher Contributions to Student Achievement by Subject, CLST Alternatives 4](#_Toc188372906)

[**Exhibit B–5.** Teacher Contributions to Student Achievement by Subject, MTEL-Flex 4](#_Toc188372907)

## 

## Executive Summary

Licensure is one of the primary mechanisms by which states ensure that new teachers have the requisite subject matter knowledge and communications and literacy skills for classroom instruction. Subject matter and basic skills tests are a key component of licensure in nearly every state (National Council on Teacher Quality, 2020). Ample evidence shows that the knowledge measured by these tests matters for student learning (Clotfelter et al., 2007, 2010; Cowan et al., 2023; Goldhaber, 2007; Goldhaber et al., 2017; Sass, 2015). However, there are also unresolved academic debates about the extent to which testing requirements exclude potentially effective teachers or limit the diversity of the workforce (Angrist & Guryan, 2004, 2008; Bennet et al., 2006; Gershenson, 2021; Gitomer et al., 2011; Rucinski & Goodman, 2019).

In 2020, the Massachusetts Board of Elementary and Secondary Education (BESE) authorized a pilot of alternatives to the Massachusetts Tests for Educator Licensure (MTEL) for educator candidates. The pilots were intended to expand access to the teaching profession, allowing prospective teachers to demonstrate their capacities in different ways while maintaining high standards for subject matter knowledge and communications and literacy skills. They were also designed to reduce barriers for teacher candidates of color, teachers already in the workforce, and out-of-state teachers to earn licensure in Massachusetts (Massachusetts Department of Elementary and Secondary Education, 2021).

The National Center for Analysis of Longitudinal Data in Education Research (CALDER) at the American Institutes for Research has engaged with the Massachusetts Department of Elementary and Secondary Education (DESE) for a multiyear evaluation of the MTEL alternative assessment pilots. In this final report, we describe participation in the alternative assessment pilots and evidence of its relationship to licensure, employment, and teacher effectiveness.

### Data Sources

The research team collected data on participation in the pilots from Spring 2021 through June 2024. As of June 30, 2024, 3,947 teacher candidates had taken at least one of the pilot assessments. This total represents about 9% of all candidates taking an MTEL during the period. In this report, we link data on these pilot participants to educator licensure and preparation program completion records through September 1, 2024 and employment records through the 2023-24 school year. The research team also linked data on pilot participants to student academic data from the 2020-21 through 2023-24 school years and performance evaluations through the 2022-23 school year. The student data include results on the Massachusetts Comprehensive Assessment System (MCAS), an index of non-tested outcomes (attendance, discipline, on-time grade progression, and grades), and the Views of Climate and Learning (VOCAL) survey of school climate.[[1]](#footnote-2) Finally, we conducted 11 interviews with faculty at the five educator preparation providers offering the attestation alternative (defined further below). These interviews covered the development of the assessment protocol, the support offered to teacher candidates, experiences operating the attestation alternative, and future plans for the attestation option if it were to continue beyond the pilot.

### Summary of Alternatives

Candidates for educator licensure typically take two types of MTEL. The *Communications and Literacy Skills Test* (CLST) is a test of basic literacy skills required of all candidates for licensure. In addition, each license field has one or more corresponding *MTEL Subject Test(s)* that focuses on the subject matter knowledge required for that field. The pilot includes four types of MTEL alternative assessments.

**CLST Alternatives:** The five approved alternatives to the CLST in this pilot are standardized assessments offered by the Evaluation Systems group of Pearson (ES) or the Educational Testing Service (ETS) for licensure in other states: Praxis Core, the National Evaluation Series (NES) Essential Academic Skills, the Missouri General Education Assessment (MoGEA), the Washington [state] Educator Skills Test-Basic (WEST-B), and the Indiana Core Academic Skills Assessment (CASA). Each of the alternatives includes a reading and writing subtest, and candidates can substitute a passing score on these tests for the CLST requirements.

**MTEL-Flex:** MTEL-Flex is offered by the Evaluation Systems group of Pearson (ES) to candidates receiving a near-passing score on one of nine MTEL subject tests: English, English as a Second Language, Foundations of Reading, General Science, General Curriculum (Subtest 1 and 2), History/Social Science, Spanish, Mathematics (Middle School), and Mathematics (Secondary). Candidates who meet the eligibility requirements may submit a written analysis of a topic covered by the relevant test rather than retake the full MTEL.

**Preparation Program Attestation:** The Preparation Program Attestation option is available for candidates enrolled in an educator preparation program at one of five organizations that have been approved based on DESE-issued guidelines to offer the attestation. Candidates take assessments and/or submit a variety of materials documenting their understanding of content knowledge for review by a content expert. The programs and tests included in the attestation option are:

* General Curriculum subtest 2 (Fitchburg State University)
* Foundations of Reading (Bridgewater State University, Fitchburg State University, Merrimack College, Westfield State University)
* English as a Second Language (Bridgewater State University, Merrimack College)
* Music MTEL (Bridgewater State University)

**Other Approved Alternatives:** DESE has approved two alternative subject matter knowledge assessments offered by ETS in English to Speakers of Other Languages and Physical Education. Based on requirements outlined by DESE, these alternative assessments are limited to candidates who have taught for at least one year under a waiver, emergency license, or temporary license; in certain non-traditional public or private schools; or as a long-term substitute or paraprofessional.

### Key Findings

#### Participation

The 2023-24 academic year saw the highest numbers of participants across all piloted assessments. Exhibit 1 describes the total cumulative participation in the pilots as of June 30, 2024. Overall, 3,947 teacher candidates have taken at least one of the pilot assessments. The CLST alternatives had the most submissions (3,698) and candidates (2,045), followed by the MTEL-Flex (2,513 submissions from 1,859 candidates). Participation was lower in the program attestation (151 candidates) and the other alternative (32 candidates) options. Some candidates have taken more than one type of alternative assessment.

Exhibit 1. MTEL Alternatives Participation

| **Alternative Assessment** | **Number of Submissions** | **Number of Candidates** |
| --- | --- | --- |
| CLST Alternatives | 3,698 | 2,045 |
| MTEL-Flex | 2,513 | 1,859 |
| Program Attestation | 154 | 151 |
| Other Approved Alternatives | 34 | 32 |

#### Teacher Background and Identity

##### Communication and Literacy Skills Tests Alternatives

Compared to teachers who passed the traditional CLST, teachers who used the CLST alternatives to meet their licensure requirements were:

* More likely to be Black or Hispanic
* More likely to have an emergency license at the time of testing
* More likely to have completed an out-of-state teacher preparation program

##### MTEL-Flex

Compared to teachers who passed the traditional MTEL subject tests, teachers who used the MTEL-Flex to meet their licensure requirements were:

* More likely to be Black
* More likely to have an emergency license at the time of testing
* More likely to be working as teachers in Massachusetts public schools at the time of testing

##### Preparation Program Attestation

The attestation option is offered at five educator preparation providers that jointly produce 23% of all program completers in Massachusetts each year. Compared to teachers from these providers passing the traditional MTELs for which their provider offer attestation, teachers who used the attestation option to meet their licensure requirements were:

* More likely to be Black or Hispanic
* More likely to have an emergency license at the time of testing

#### Access to the Profession

##### Communication and Literacy Skills Tests Alternatives

The CLST alternatives are generally less expensive than the traditional CLST MTEL. In a survey of participants, 95% reported that ease of accessing the test was somewhat or very important in their decision to take an alternative, while 85% reported the same for the cost.

Adjusting for candidates’ prior testing performance, overall pass rates across all CLST alternatives are slightly higher than for the traditional MTEL. However, pass rates vary considerably across the test series.

Across CLST alternatives, candidates obtained licensure and employment in Massachusetts public schools at similar rates to those taking the traditional tests.

##### MTEL-Flex

Candidates taking the MTEL-Flex were about 16 percentage points more likely to pass the test than candidates who qualified for MTEL-Flex but chose to retake the traditional MTEL. There has been some variation in pass rates on the MTEL-Flex across the pilot period: After declining following its introduction, pass rates on the MTEL-Flex increased during the 2022-23 and 2023-24 school years.

Teachers who took the MTEL-Flex had higher licensure and employment rates by about 2-3 percentage points compared to candidates who qualified for MTEL-Flex but chose to retake the traditional MTEL.

In addition to higher pass rates, the MTEL-Flex also appears to have increased retesting rates for candidates who qualified after initially failing the MTEL. The option to take the MTEL-Flex increased retesting rates by about 3 percentage points in this group.

##### Preparation Program Attestation

Candidates participating in the attestation option were about 30 percentage points more likely to pass the MTEL than a comparison group of other candidates enrolled in an educator preparation program offered by their provider who took the traditional MTEL.

They were about 13 percentage points more likely to complete their preparation program and 8 percentage points more likely to find employment as teachers.

#### Teacher Effectiveness

##### Communication and Literacy Skills Tests Alternatives

Teachers who passed a CLST alternative performed similarly on teacher evaluations and had similar contributions to overall student achievement as teachers passing the traditional CLST MTEL. They may have been less effective in math instruction and at improving a battery of non-tested outcomes than teachers passing the CLST MTEL, but these differences were not statistically significant at conventional levels.

##### MTEL-Flex

Teachers who passed an MTEL-Flex made similar contributions to student growth on MCAS, non-tested outcomes, and school climate as teachers passing the traditional MTEL subject tests.

Teachers who passed the MTEL-Flex were less likely to earn exemplary ratings (the highest performance category in Massachusetts) than candidates passing the traditional MTEL subject tests, but they were no more likely to earn low performance evaluations (unsatisfactory or needs improvement).

##### Preparation Program Attestation

Teachers whose educator preparation program attested to their subject matter knowledge received similar performance evaluations as candidates who passed the traditional MTEL subject tests. However, due to the small number of participants in this option, we cannot rule out the possibility that there were meaningful differences in performance evaluations. There were also too few teachers participating in this route that could be linked to their students to assess their effects on student outcomes.

## Introduction and Background

Massachusetts statutorily requires all candidates for licensed educator roles to pass a series of tests before they obtain a license.[[2]](#footnote-3) These include the Communication and Literacy Skills Test (CLST), a two-part test that covers reading comprehension and writing proficiency that is required for all educator licenses. In addition, teacher candidates must pass subject tests aligned with the [state curricular frameworks](https://www.doe.mass.edu/frameworks/current.html) and [Subject Matter Knowledge Guidelines](https://www.doe.mass.edu/edprep/domains/instruction/smk-guidelines.docx) in the field in which they intend to teach. The tests are designed to ensure that candidates demonstrate a functional level of understanding of the subject matter knowledge and communications and literacy skills required for the license.

Teacher content knowledge contributes to student learning. Researchers have demonstrated in a variety of contexts that students who are assigned to teachers with stronger basic academic skills and higher mastery of their subject area content, as measured by licensure tests, perform better on standardized assessments (Clotfelter et al., 2007, 2010; Cowan et al., 2023; Goldhaber, 2007; Goldhaber et al., 2017; Sass, 2015). At the same time, testing requirements may exclude some teachers who would become effective educators (Gershenson, 2021). Passing rates on licensure tests also tend to be lower for teacher candidates of color, raising concerns that testing requirements may limit the diversity of the teacher workforce (Nettles et al., 2011; Rucinski & Goodman, 2019).

The Massachusetts Board of Elementary and Secondary Education (BESE) authorized the Pilot of Alternative Assessments to the MTEL on October 20, 2020, to identify and study alternative mechanisms for assessing communications and literacy skills and subject matter knowledge. The alternative assessments are intended to be more inclusive in their approach to identifying strong educators while maintaining high standards for the profession (Department of Elementary and Secondary Education [DESE], 2021). There are currently four types of alternative assessment models included in the pilot:

**CLST Alternatives:** The five approved alternatives to the CLST in this pilot are standardized assessments offered by the Evaluation Systems group of Pearson (ES) or the Educational Testing Service (ETS) for licensure in other states: Praxis Core, the National Evaluation Series (NES) Essential Academic Skills, the Missouri General Education Assessment (MoGEA), the Washington [state] Educator Skills Test-Basic (WEST-B), and the Indiana Core Academic Skills Assessment (CASA). Each of the alternatives includes a reading and writing subtest, and candidates can substitute a passing score on these tests for the CLST requirements.

**MTEL-Flex:** MTEL-Flex is offered by the Evaluation Systems group of Pearson (ES) to candidates receiving a near-passing score on one of nine MTEL subject tests: English, English as a Second Language, Foundations of Reading, General Science, General Curriculum (Subtest 1 and 2), History/Social Science, Spanish, Mathematics (Middle School), and Mathematics (Secondary). Candidates who meet the eligibility requirements may submit a written analysis of a topic covered by the relevant test rather than retake the full MTEL test.

**Preparation Program Attestation:** The Preparation Program Attestation option is currently approved for candidates enrolled in an educator preparation program at one of five organizations that have been approved to offer the attestation by the Department. Candidates take assessments and/or submit a variety of materials documenting their understanding of content knowledge for review by a content expert. The providers and tests included in the pilot are:

* General Curriculum subtest 2 (Fitchburg State University)
* Foundations of Reading (Bridgewater State University, Fitchburg State University, Merrimack College, Westfield State University)
* English as a Second Language (Bridgewater State University, Merrimack College)
* Music MTEL (Bridgewater State University)

**Other Approved Alternatives:** DESE has approved two alternative subject matter knowledge assessments offered by ETS in English to Speakers of Other Languages and Physical Education. Based on requirements outlined by DESE, these alternative assessments are limited to candidates who have taught for at least one year under a waiver, emergency, or temporary license; in certain non-traditional public or private schools; or as a long-term substitute or paraprofessional.

Prior annual reports assessed the introduction of the alternatives, participant beliefs and experiences, and limited outcomes through June 2023. In this update, we include data on pilot participants over the past year and match them to more recent data on licensure, employment performance evaluations, and student outcomes. We also report on interviews conducted with faculty at each of the preparation providers offering the attestation option.

## Candidate Experiences

|  |
| --- |
| Communication and Literacy Skills Tests Alternatives By June 30, 2024, 2,045 candidates had taken an approved CLST alternative assessment. The most popular alternative assessments were reading and writing tests accepted for licensure in Indiana (CASA) and Missouri (MoGEA). The two national assessments (NES and Praxis) approved through the pilot were less popular.  The approved CLST alternatives are generally less expensive than the CLST MTEL. In a survey of participants, most respondents reported that ease of accessing the test (95%) or the cost of the test (84%) were somewhat or very important in their decision to take an alternative. MTEL-Flex By June 30, 2024, 1,859 candidates had taken an MTEL-Flex alternative assessment. The most popular MTEL-Flex assessments were those associated with the elementary license, General Curriculum and Foundations of Reading. Preparation Program Attestation By June 30, 2024, 151 candidates across five institutions had participated in an approved program attestation.  Across educator preparation programs and content areas, significant variation existed regarding the structure of the assessment, the mode in which students demonstrated mastery, how students were supported, and the criteria for participation.  Preparation providers believe their alternative assessment either meets or exceeds the rigor of the traditional MTEL. In addition to describing the questions and assignments as difficult, they also stated that the students must apply their knowledge in ways that teachers do, which is asking candidates to be more fluent with content knowledge than they may need to be for a multiple-choice exam. |

### Communications and Literacy Skills (CLST) Alternatives

The CLST MTEL is a two-part test in reading and writing required of all candidates in Massachusetts, including those seeking administrative, specialist, or professional support licenses. Candidates must earn a passing score on each of the two subtests to advance to licensure. As part of the pilot, DESE accepted similar assessments offered in other states to satisfy the CLST requirements. DESE has approved five alternatives to the CLST: Praxis Core, the National Evaluation Series (NES) Essential Academic Skills, the Missouri General Education Assessment (MoGEA), the Washington [state] Educator Skills Test-Basic (WEST-B), and the Indiana Core Academic Skills Assessment (CASA). Each of these tests includes a reading and writing subtest, and candidates could submit a passing score on either subtest to meet their licensure requirements. In Exhibit 2, we plot the cumulative number of CLST submissions over the course of the pilot. By June 30, 2024, 2,045 candidates had completed 3,698 subtest submissions.

Exhibit 2. Cumulative CLST Alternative Submissions by Date

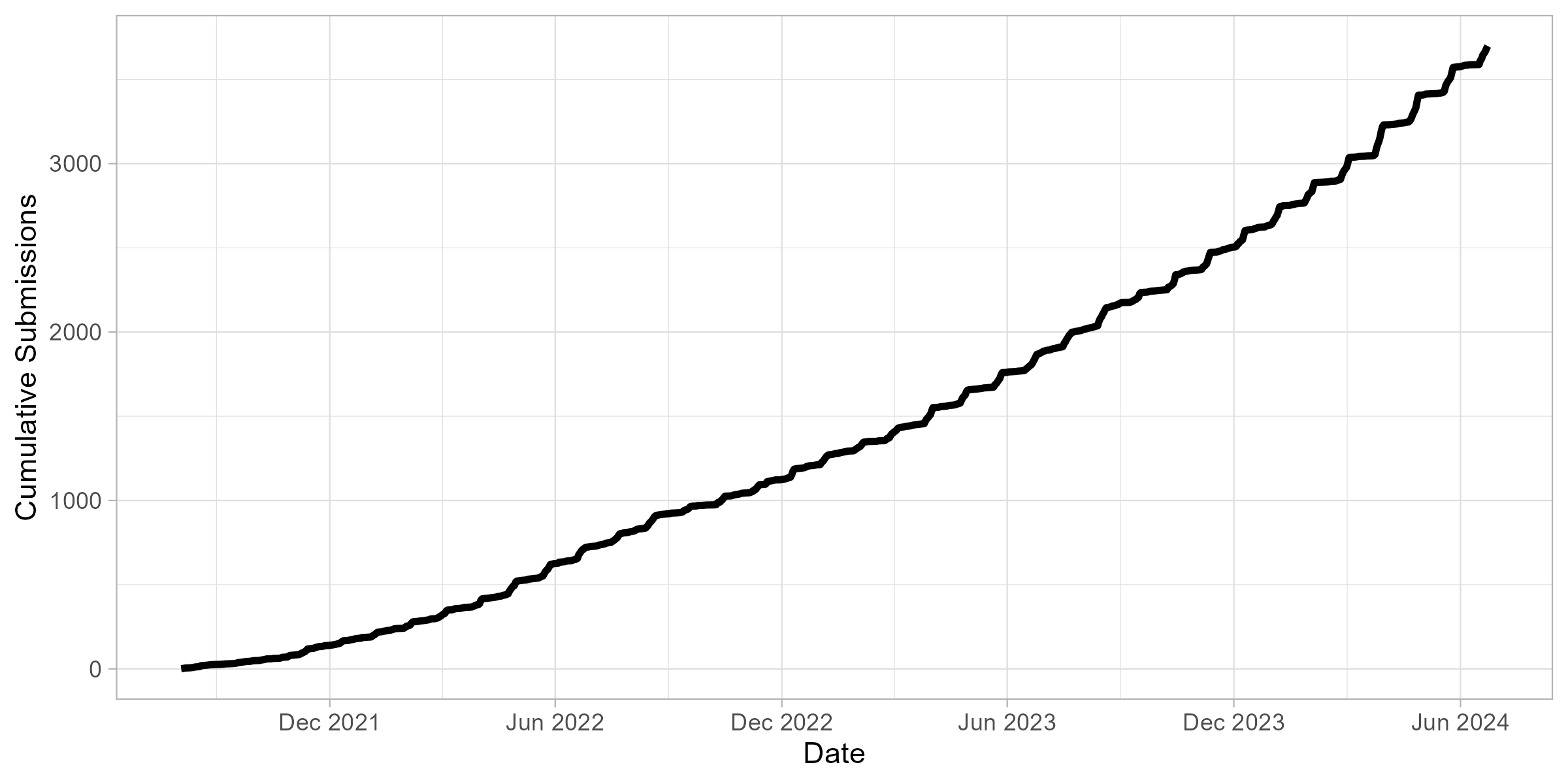
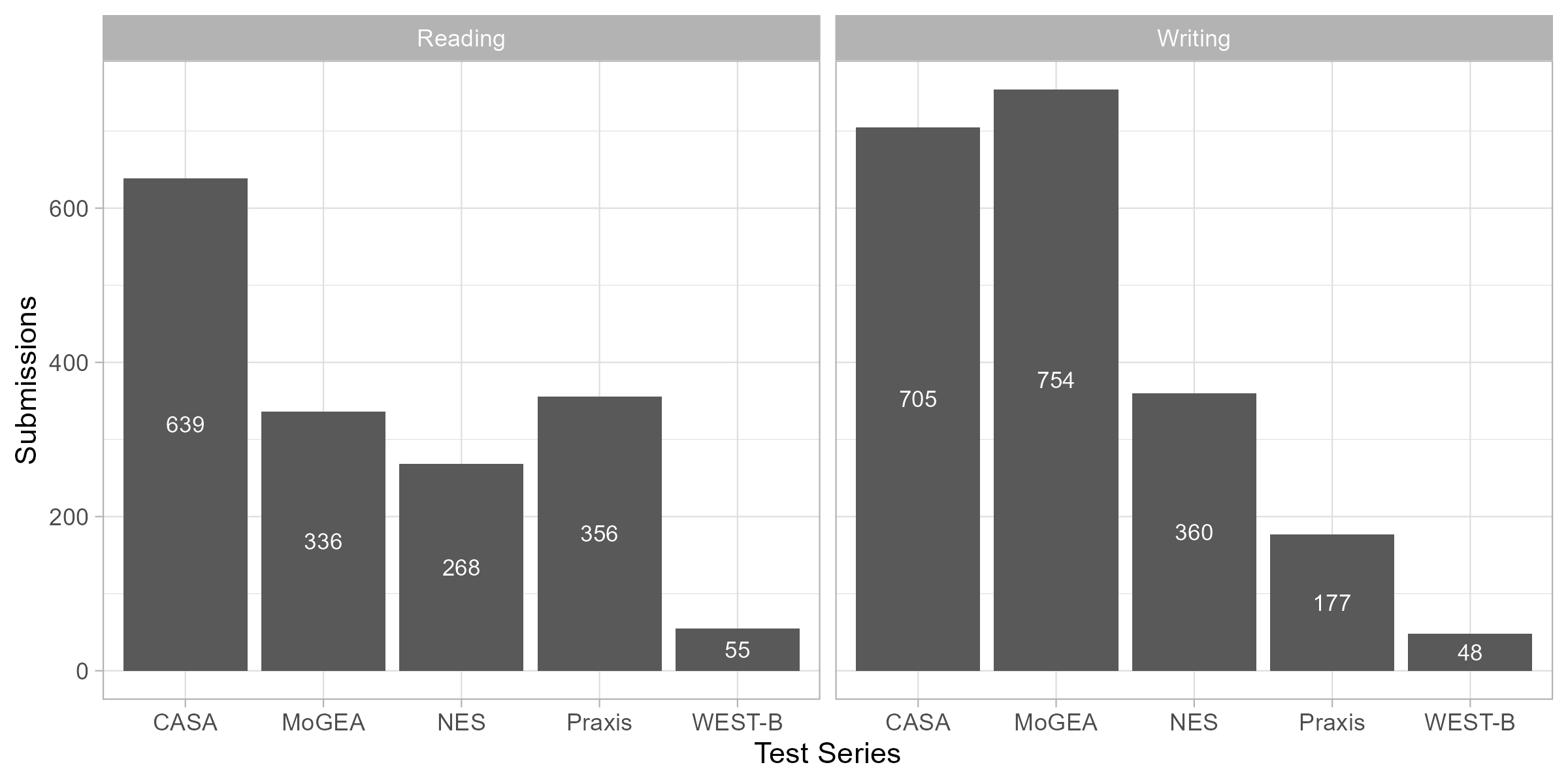


Exhibit 3. CLST Alternative Submissions by Test Series and Subtest



In Exhibit 3, we show participation by test series. The CASA and MoGEA were the most popular CLST alternatives, followed by the NES and Praxis. Few candidates have taken the WEST-B. The Praxis and NES are accepted in several other states, while the CASA and MoGEA are offered in Indiana and Missouri, respectively.

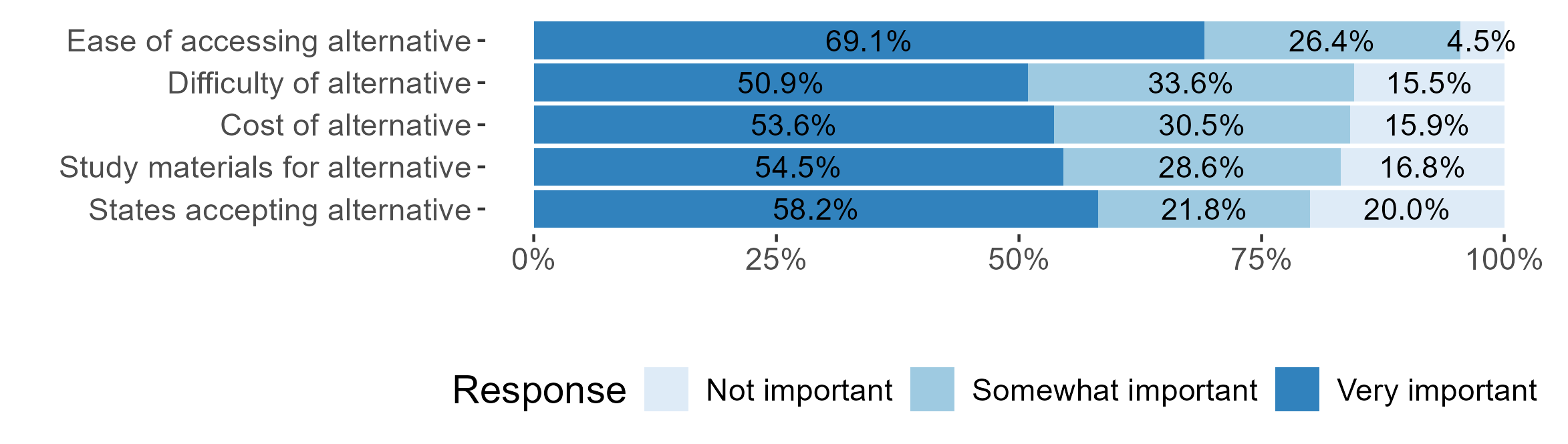
Given the small number of states accepting the most popular CLST alternatives tests, we briefly highlight other features of the tests in Exhibit 4. The CLST MTEL and Praxis are the most expensive options at $112 and $110, respectively. Although the NES is less expensive ($75), it is comparable in price to the CASA ($76) and MoGEA ($50). The CASA and MoGEA also have fewer open-response items (1) than both the MTEL and Praxis (2). Finally, the CASA, MoGEA, and WEST-B are also offered only in an online format. The lower cost and online-only proctoring options may explain part of their popularity.

Exhibit 4. Cost and Format of CLST MTEL and Communications and Literacy Skills Alternatives

|  | **Maximum Time Allowed for Reading/Writing** | **Test Structure** | **Cost per Subtest (Combined)** | **Online** |
| --- | --- | --- | --- | --- |
| MTEL | 240/240 minutes | Reading: 42 multiple-choice questions  Writing: 35 multiple-choice questions, 7 sentence correction items, 2 open-response items | $69.50 ($112) | Online proctoring available |
| CASA | 75/105 minutes | Reading: 40 multiple-choice questions  Writing: 42 multiple-choice questions and 1 constructed-response question | $38 ($76) | Online proctoring only |
| MoGEA | 75/60 minutes | Reading: 39 multiple choice questions  Writing: 1 written assignment | $25 ($50) | Online proctoring only |
| NES | 60/75 minutes | Reading: Approximately 45 multiple-choice questions  Writing: Approximately 46 multiple-choice items, 1 written assignment | $50  ($75) | Online proctoring available |
| Praxis | 100/75 minutes | Reading: 56 selected-response questions  Writing: 40 selected-response questions and two essay questions | $90  ($110) | Online proctoring available |
| WEST-B | 120/180 minutes | Reading: 60 multiple choice questions  Writing: 50 multiple choice questions and 2 constructed-response questions | $32 ($64) | Online proctoring only |

*Notes:* Characteristics of the MTEL and CLST Alternatives. Length indicates the length of the reading/writing subtests, respectively. Test structure describes the item count and type for reading and writing subtests. Cost indicates the cost of each subtest. The cost in parentheses indicates the cost of registering for both subtests simultaneously. Online indicates that there is an online testing option.

Exhibit 5. Importance of Factors Influencing Decision to Take CLST Alternatives



*Source:* Survey of teacher candidates who took at least one CLST alternative during or before academic year 2022-2023.

We surveyed participants in summer 2023 about their decision to take a CLST alternative. The 2023 Annual Report describes the survey findings in more detail, but candidates’ reported reasons for taking the CLST alternatives (Exhibit 5) appear to reflect some of the differences between the tests noted in Exhibit 4. Among respondents, 95% reported that the ease of accessing the test was somewhat or very important in their decision to take the alternative; 85% reported that the difficulty was either somewhat or very important; and 84% reported that the cost was somewhat or very important. Although 58% of respondents reported that the number of states accepting the alternative assessment was very important in their decision, only 80% reported that it was either somewhat or very important. The CLST alternative was intended to provide additional access to licensure for out-of-state candidates, but it appears that candidates also had other reasons for taking these assessments.

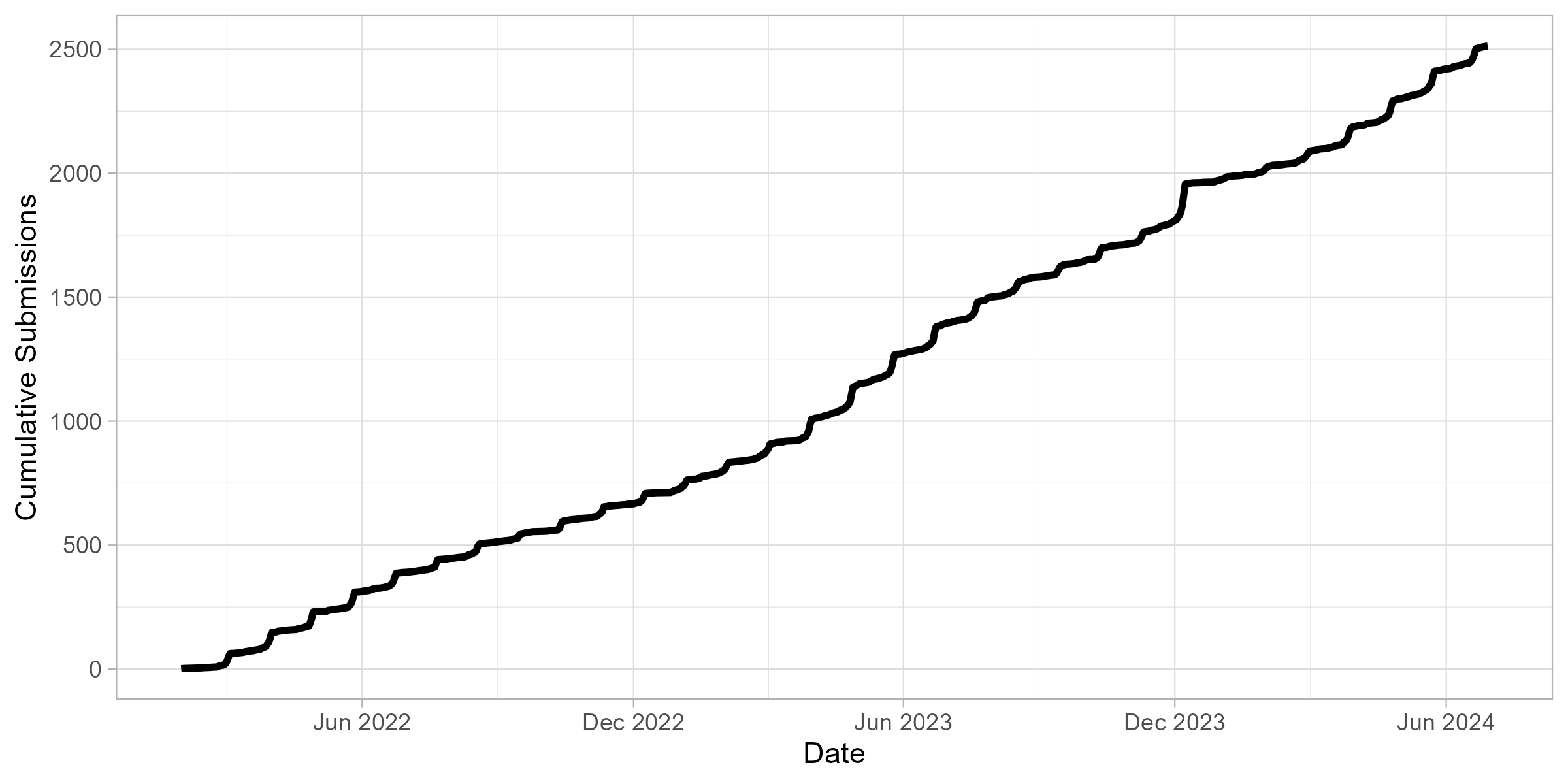
### MTEL-Flex

The MTEL-Flex is an alternative for candidates receiving a near-passing score on one of nine MTEL subject tests: English, English as a Second Language, Foundations of Reading, General Science, General Curriculum (Subtest 1 and 2), History/Social Science, Spanish, Mathematics (Middle School), and Mathematics (Secondary). To qualify for the MTEL-Flex, candidates must have taken the traditional MTEL and received a score less than one standard error of measurement (SEM) below the minimum passing score, which varies by test.[[3]](#footnote-4) In the 2022 and 2023 Annual Reports, we discussed the eligible pool of candidates and noted that, in some cases, as many as 20% of all submissions on the traditional MTEL qualified for MTEL-Flex. Qualification for MTEL-Flex is automatic, and eligible candidates have the option to register when they log into the ES website to schedule an MTEL retake (DESE, 2022). The DESE and Pearson websites also provide information about the MTEL-Flex assessment.

The MTEL-Flex varies in several important respects from the traditional MTEL. The traditional MTEL includes some open response questions, but most of the points come from multiple choice questions. By contrast, MTEL-Flex consists of a three-page written analysis of a topic from the relevant curricular frameworks. Candidates complete the submission in their own time and can consult primary sources (e.g., textbooks) as they prepare their submissions. The MTEL-Flex assessment is less expensive ($49–$69) than retaking the corresponding MTEL subject tests ($94–$139). Candidates who fail the MTEL-Flex assessment are provided a scoring report with feedback and may resubmit the test (DESE, 2022). Candidates who receive a passing score are considered to have completed their licensure testing requirement.

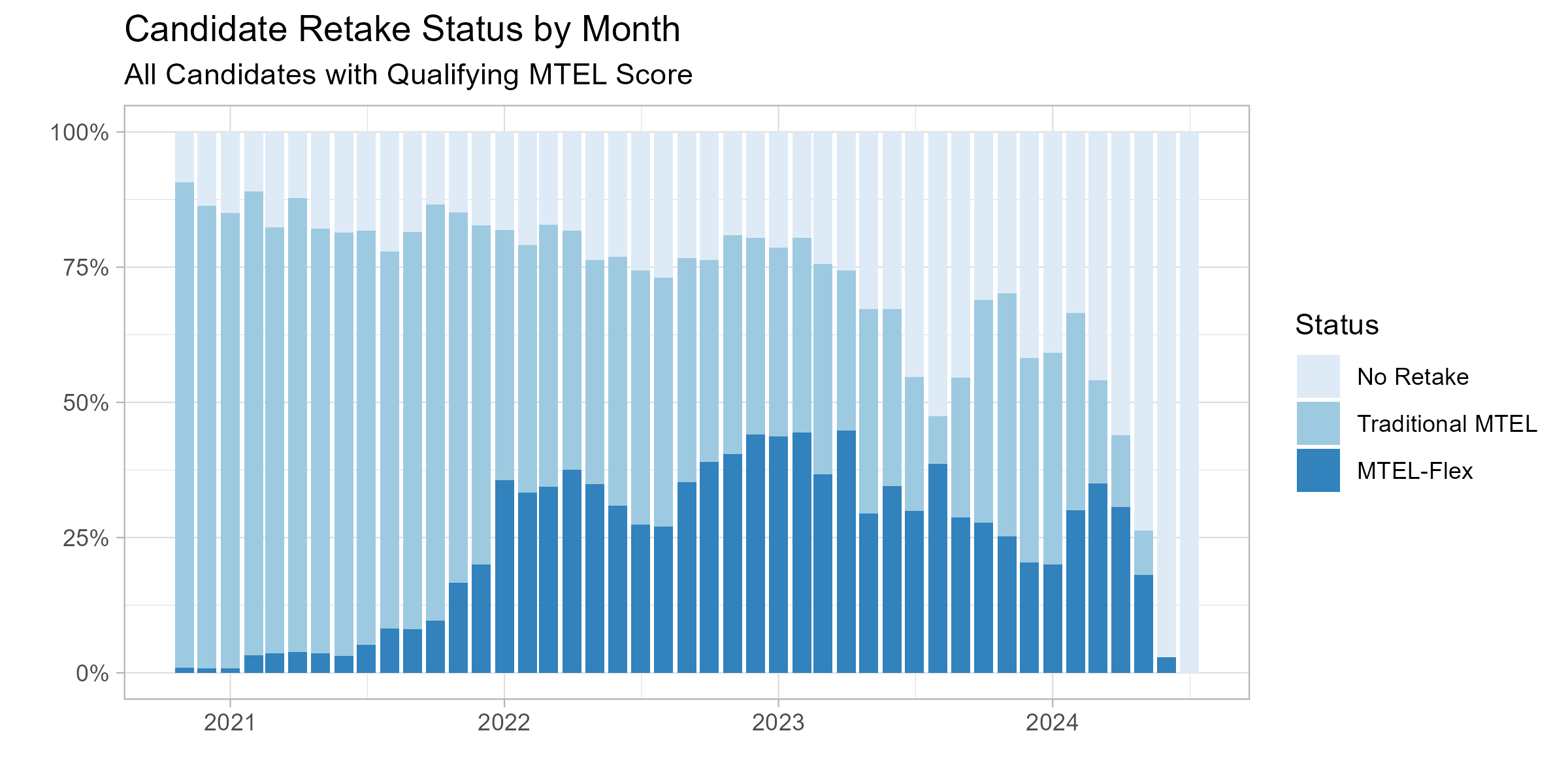
Since the 2023 Annual Report, the number of total MTEL-Flex submissions has approximately doubled to 2,513 (Exhibit 6).

Exhibit 6. Cumulative MTEL-Flex Submissions by Date



In Exhibit 7, we plot the retesting decision for each candidate qualifying for the MTEL-Flex since the start of the pilot. Since 2022 (when the MTEL-Flex first became available), about 24% of qualifying candidates have chosen to take the MTEL-Flex. By comparison, 46% chose to reattempt the traditional MTEL, and about 30% chose not to reattempt either version of the test. The decline in MTEL-Flex participation during 2023 and 2024 coincides with the introduction of several new MTELs; the MTEL-Flex did not become available for the new tests until several months after their introduction. The decline in the rate of candidates retaking any test after 2024 reflects the fact that many candidates spend multiple months preparing between attempts. Hence, some of these candidates will likely reattempt the MTEL in the future.

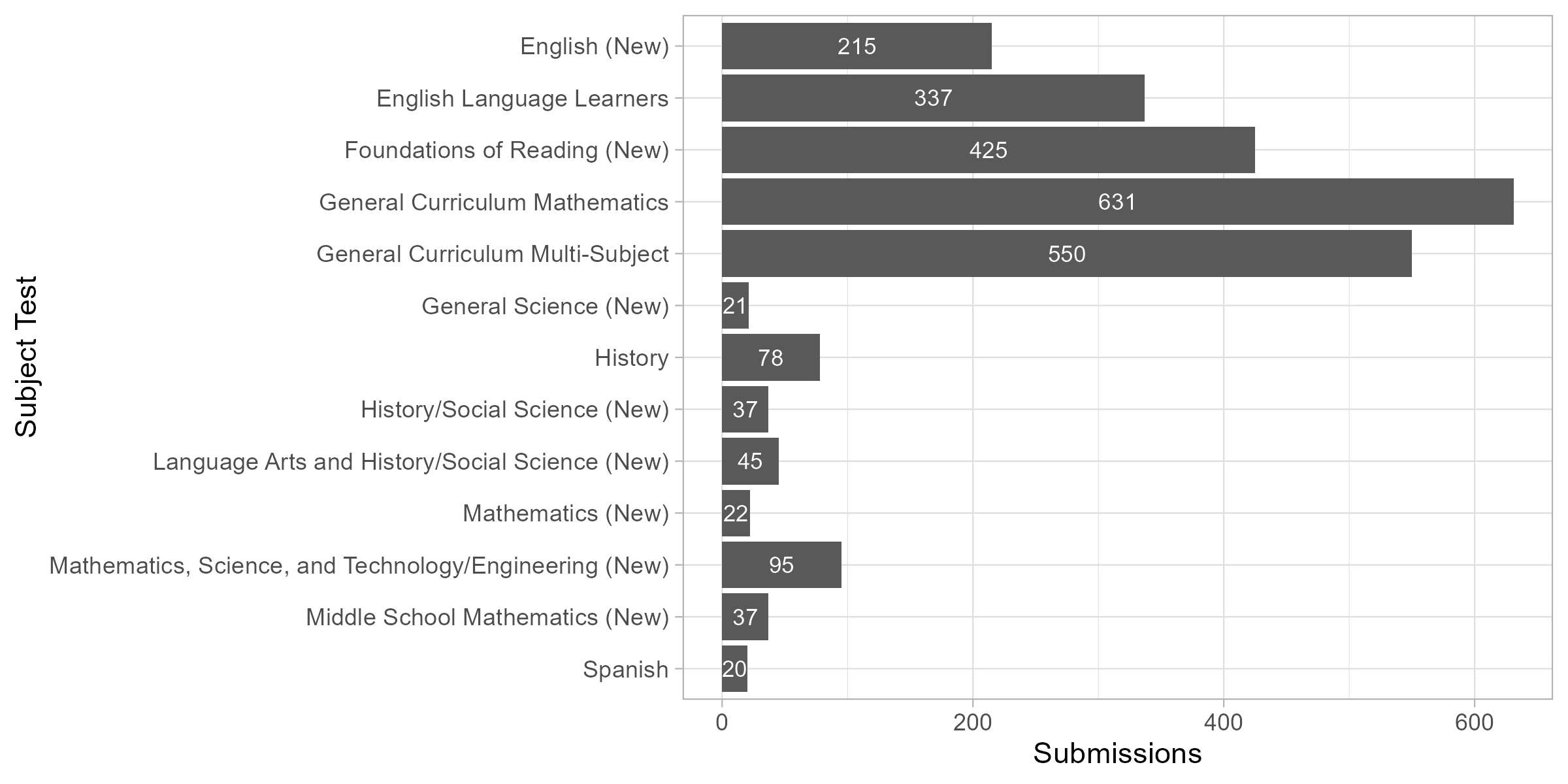
Exhibit 7. Candidate MTEL-Flex Retake Status by Month of Qualification



*Notes:* Retest decisions for candidates qualifying MTEL-Flex. The sample includes all teacher candidates who have received a score on the subject MTEL that qualify them for the MTEL-Flex. “No Retake” includes candidates who do not retake either the traditional MTEL or MTEL-Flex before June 30, 2024.

As indicated in Exhibit 8, the General Curriculum tests (Mathematics; Multisubject; Mathematics, Science, and Technology/Engineering; Language Arts and History/Social Science) and the Foundations of Reading test, frequently taken by candidates seeking elementary licensure, were the most popular MTEL-Flex tests. The newer version of the General Curriculum test (Language Arts and History/Social Science and Mathematics, Science, and Technology/Engineering) has only been available as an MTEL-Flex option since March 2024, so relatively few candidates took this test prior to June 30, 2024. The next most common tests are the English Language Learners and English tests, which have been available since the beginning of the pilot. Each of the remaining tests has fewer than 100 submissions.

Exhibit 8. MTEL-Flex Submissions by Subject Test

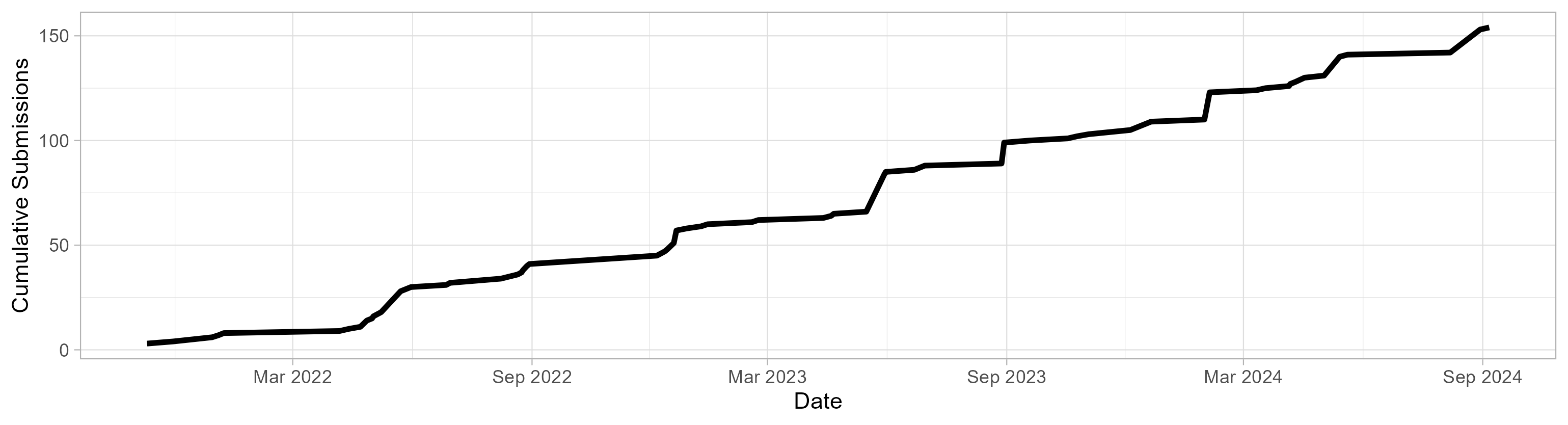


*Notes:* Total MTEL-Flex submissions by MTEL subject test. Tests marked as (New) are redesigned versions of the MTEL that have been introduced during the pilot. The Mathematics, Science, and Technology/Engineering and Language Arts and History/Social Science tests were introduced during the pilot to collectively replace the General Curriculum Mathematics and Multisubject Tests.

### Program Attestation

The Preparation Program Attestation option is an assessment offered by five approved educator preparation providers to document teacher candidates’ content area knowledge: Bridgewater State University, Fitchburg State University, Merrimack College, Westfield State University, and Worcester State University. DESE released [guidelines](https://www.doe.mass.edu/edprep/resources/guidelines-advisories/alt-assess-guide.pdf) for program attestations in February 2021 and the first candidates completed the process in late 2021. We plot the cumulative participation in Exhibit 9. By June 30, 2024, 151 candidates had participated in an approved program attestation.

Exhibit 9. Cumulative Program Attestations by Date



We conducted interviews with 11 faculty who are either program coordinators or content experts responsible for designing and/or administering the alternative assessments. In this section, we provide a high-level summary of their perceptions about the process of developing and implementing the attestation and their interest in continuing and potentially expanding them. We provide a more complete discussion of the findings from the interviews in Appendix A.

Most faculty reported the development process was driven by individual faculty or several faculty after the call for applications was issued. To date, only one program has developed an attestation for the General Curriculum Test, which is unique in that it crosses multiple departments within the educator preparation program and requires the content expertise of multiple academic faculty. Respondents reported that this attestation required additional coordination and necessitated multiple people working together across departments, making it more complicated than other attestations to both develop and administer.

The structure of the program attestation varies across programs and subject areas. In three educator preparation programs, students only demonstrate mastery in the subareas in which they did not have an adequate score on the traditional MTEL; at two programs, students must demonstrate mastery in all objectives of the MTEL. Students typically can resubmit materials or retake an attestation test if they do not pass the first time. Preparation for and support of students seeking to take the program attestation varies across programs and subject tests. Several programs offer courses to help students prepare for the assessment, although program administrators also reported spending a lot of time providing individualized support for each student beyond any required or optional coursework. Faculty reported students typically spend several months preparing for the attestation. All respondents reported that they believe their alternative assessment either meets or exceeds the expectations and rigor of the traditional MTEL.

Respondents shared mixed responses about whether the program attestation has helped expand access to the profession. Several respondents, however, reported a belief that the ESL alternative assessment has helped to diversify the workforce in terms of increasing the number of teachers of color as well as the number of nontraditional candidates, in terms of age and people whose first language is not English. Several respondents reported that it took a little bit of time for students to become aware of the alternative assessment opportunity, despite their efforts to get the word out. Multiple people said that interest has been growing as information about the opportunity has spread and that more students are now aware of the alternative assessment.

Respondents volunteered different aspects of the attestation that they believe are working well. Multiple respondents reported that the program attestation has enabled people who they believe are and will be strong teachers to get licensed. Multiple respondents also reported that providing the alternative assessment free-of-charge to the teacher candidates is an important advantage, making it more equitable and financially accessible to those candidates who are already enrolled in an educator preparation program.

Faculty did describe some challenges with the attestation, although they were positive about continuing the option. The three main challenges that respondents identified are the lack of funding to support this initiative; some of the bureaucratic requirements, either getting past assignments from other professors or in uploading the data to the Department of Elementary and Secondary Education (DESE); and getting the program up and running. At three of the educator preparation providers, the content experts are running the program with no extra pay or release from teaching. All 11 people interviewed enthusiastically stated interest in continuing with the program attestation.

## Candidate Background and Identity

|  |
| --- |
| Key findings Communication and Literacy Skills Tests Alternatives Candidates using the CLST alternatives to meet licensure requirements were disproportionately likely to be teachers working in Massachusetts public schools on emergency licenses; to have completed out-of-state teacher preparation programs; and to be Black or Hispanic educators. MTEL-Flex Candidates using the MTEL-Flex to meet licensure requirements were disproportionately likely to be teachers working in Massachusetts public schools on emergency licenses; and to be Black educators. Preparation Program Attestation Candidates using the program attestation option to meet licensure requirements were disproportionately likely to be teachers working in Massachusetts public schools on emergency licenses; and to be Black or Hispanic educators. |

In this section, we discuss the background of candidates participating in the MTEL alternatives pilot. We consider the ethnoracial identities of candidates, as well as the diversity of licensure pathway, employment status, and preparation background.

### Communications and Literacy Skills Alternatives

We present descriptive statistics for candidates who took either a CLST alternative or the traditional CLST after October 20, 2020 (Exhibits 10 and 11). In the first two columns, we show characteristics for all candidates who submitted either the traditional or alternative CLSTs. In the next two columns, we show characteristics of candidates who received passing scores on the traditional MTEL CLST and the CLST alternatives. This latter group consists of the candidates who have used either option to satisfy their licensure requirements in Massachusetts.

Candidates taking a CLST alternative were more likely to be working as teachers on an emergency license than candidates taking the traditional CLST. About 36% of those taking the alternatives were on an emergency license; this was true for only 21% of candidates taking the traditional MTEL. Similarly, 23% of CLST alternative test takers and 16% of traditional test takers were employed as teachers when they took the test. Out-of-state candidates were also more likely to take a CLST alternative. Among candidates whom we linked to licensure records, 17% of those taking an alternative and 13% of those taking the traditional MTEL graduated from an out-of-state program.[[4]](#footnote-5) The same patterns hold when we look only at candidates passing the test.

Exhibit 10. Teacher Candidate Characteristics and CLST Alternatives Choice

|  | All Submissions | | Passing Submissions | |
| --- | --- | --- | --- | --- |
|  | % Traditional MTEL | % CLST Alternative | % Traditional MTEL | % CLST Alternative |
| Emergency License | 21.1 | 35.6 | 18.8 | 28.3 |
| Out of State | 12.8 | 16.9 | 13.4 | 18.1 |
| Prior Program Completion | 2.7 | 4.2 | 2.7 | 3.7 |
| Current Teacher | 15.6 | 22.7 | 14.0 | 18.2 |
| Current Other Educator | 6.3 | 8.4 | 6.3 | 7.7 |
| Test Attempt | 1.4 | 2.5 | 1.2 | 1.8 |

*Notes:* Average characteristics of traditional and alternative MTEL test-takers. Sample includes all submissions on traditional or alternative CLSTs after October 20, 2020. Prior program enrollment, prior program completion, and emergency license are all measured as of the testing date. Employment outcomes are measured during the academic year a candidate took the test.

Exhibit 11. Teacher Candidate Characteristics and CLST Alternatives Choice

|  | All Submissions | | | Passing Submissions | |
| --- | --- | --- | --- | --- | --- |
| Teacher Race/Ethnicity | % Traditional MTEL | % CLST Alternatives | % Traditional MTEL | | % CLST Alternatives | |
| No Response | 1.7 | 1.0 | 2.0 | | 1.5 | |
| Black | 8.0 | 13.4 | 5.4 | | 8.6 | |
| Asian or Pacific Islander | 4.6 | 4.6 | 4.7 | | 4.7 | |
| Hispanic | 10.9 | 17.0 | 8.5 | | 13.1 | |
| White | 73.0 | 61.8 | 77.7 | | 69.9 | |
| Other | 1.6 | 2.2 | 1.5 | | 2.1 | |

*Notes:* Average characteristics of traditional and alternative MTEL test-takers. Sample includes all submissions on traditional or alternative CLSTs after October 20, 2020.

Candidates taking a CLST alternative were also more ethnoracially diverse than candidates taking the traditional MTEL (Exhibit 11). About 13% of candidates were Black (compared to 8% of those who took the traditional MTEL) and 17% were Hispanic (compared to 11% of those who took the traditional MTEL).

### MTEL-Flex

We present descriptive statistics for candidates taking MTEL-Flex in Exhibits 12 and 13. The sample includes all candidates who submitted an MTEL subject test after the earliest eligibility date. The first two columns include all submissions on the MTEL-Flex and the traditional MTEL; the second two columns compare candidates who received passing scores and used either the traditional MTEL or MTEL-Flex to satisfy their licensure requirements. In all cases, we only consider MTEL subject tests where there was an available MTEL-Flex option.

Candidates taking the MTEL-Flex were more likely to be in the workforce than those taking the traditional MTEL. Among those taking the MTEL-Flex, about 40% had an emergency license, 28% had previously completed an educator preparation program, and 42% were working as teachers in Massachusetts public schools. By contrast, only about 30% of teachers taking the traditional MTEL had an emergency license or were working as teachers when they took the test. The patterns are generally similar when we compare candidates satisfying licensure requirements by passing the MTEL-Flex to those using the traditional MTEL.

Exhibit 12. Teacher Candidate Characteristics and MTEL-Flex Choice

|  | All Submissions | | Passing Submissions | |
| --- | --- | --- | --- | --- |
|  | % Traditional MTEL | % MTEL-Flex | % Traditional MTEL | % MTEL-Flex |
| Emergency License | 29.3 | 40.1 | 23.9 | 40.4 |
| Out of State | 13.4 | 8.1 | 15.1 | 8.2 |
| Prior Program Completion | 18.2 | 28.2 | 17.8 | 28.9 |
| Current Teacher | 32.9 | 42.1 | 29.0 | 44.0 |
| Current Other Educator | 4.8 | 5.3 | 4.8 | 5.0 |
| Test Attempt | 1.5 | 3.3 | 1.2 | 3.2 |

*Notes:* Average characteristics of MTEL test-takers participation in MTEL-Flex. Prior program completion and emergency license are measured as of the testing date.

Exhibit 13. Teacher Candidate Characteristics and MTEL-Flex Choice

|  | All Submissions | | | Passing Submissions | |
| --- | --- | --- | --- | --- | --- |
| Teacher Race/Ethnicity | % Traditional MTEL | % MTEL-Flex | % Traditional MTEL | | % MTEL-Flex |
| No Response | 0.6 | 0.1 | 0.8 | | 0.0 |
| Black | 5.5 | 5.1 | 3.1 | | 4.7 |
| Asian or Pacific Islander | 3.7 | 3.3 | 4.1 | | 3.9 |
| Hispanic | 8.2 | 5.0 | 7.0 | | 5.7 |
| White | 80.4 | 81.7 | 83.3 | | 80.9 |
| Other | 1.4 | 4.5 | 1.6 | | 4.4 |
| All | 100.0 | 100.0 | 100.0 | | 100.0 |

*Notes:* Average characteristics of MTEL test-takers by participation in MTEL-Flex.

Although candidates who took the MTEL-Flex were demographically similar to those taking the traditional MTEL, the pool of candidates using the MTEL-Flex to satisfy licensure requirements was more ethnoracially diverse than those passing the traditional MTEL (see Exhibit 13). Candidates *submitting* MTEL-Flex assessments were somewhat less likely to be Hispanic and somewhat more likely to report an “other” racial/ethnic identity than those taking the traditional MTEL. However, because the threshold score for qualification for the MTEL-Flex was near the median test score, the pool of potential candidates for the MTEL-Flex was somewhat less diverse than the larger pool of candidates who failed the assessments and somewhat more diverse than the pool of candidates who received passing scores on their first attempt. Thus, the patterns were slightly different when we instead compare candidates who used the MTEL-Flex to fulfil licensure requirements to those using the traditional MTEL. Those *passing* the MTEL-Flex were somewhat more likely to be Black (4.7% compared to 3.1%) or to report an “other” race/ethnicity (4.4% compared to 1.6%).

### Program Attestation

In Exhibits 14 and 15, we describe participants in the program attestation option. We compare candidates participating in the program attestation option with other candidates taking the same MTEL assessments and reporting scores to the same educator preparation programs. Among these candidates, participants in the program attestation option were more likely to hold an emergency license (31% compared to 26%), although they were slightly less likely to be employed in a teaching role during the school year in which they participated (28% compared to 29%).

Exhibit 14. Program Attestation Candidate Background

|  | % Traditional MTEL | % Attestation |
| --- | --- | --- |
| Emergency License | 25.7 | 31.2 |
| Current Teacher | 28.7 | 27.9 |
| Current Other Educator | 3.7 | 1.3 |
| Test Attempt | 2.1 | 3.0 |

*Notes:* Average characteristics of MTEL test-takers by participation in program attestation.

Exhibit 15. Program Attestation Candidate Ethnoracial Identity

|  | All Submissions | | | Passing Submissions | |
| --- | --- | --- | --- | --- | --- |
| Teacher Race/Ethnicity | % Traditional MTEL | % Attestation | % Traditional MTEL | | % Attestation |
| No Response | 12.0 | 7.1 | 13.3 | | 7.6 |
| Black | 3.8 | 7.8 | 1.8 | | 8.4 |
| Asian or Pacific Islander | 2.0 | 2.6 | 1.8 | | 1.7 |
| Hispanic | 6.2 | 14.3 | 3.7 | | 10.9 |
| White | 75.2 | 67.5 | 78.2 | | 70.6 |
| Other | 0.7 | 0.6 | 1.2 | | 0.8 |
| All | 100.0 | 100.0 | 100.0 | | 100.0 |

*Notes:* Average characteristics of MTEL test-takers by participation in program attestation.

Candidates who participated in the program attestation option were more ethnoracially diverse than other candidates enrolled in an educator preparation program offered by the same provider. About 8% of candidates participating in the attestation option were Black and about 14% were Hispanic, compared to 4% and 6%, respectively, taking the traditional MTEL. These differences are even more pronounced when looking at passing submissions.

## Access to the Profession

|  |
| --- |
| Key findings Communication and Literacy Skills Tests Alternatives Pass rates on the CLST alternatives varied significantly across the test series, with the CASA (Indiana) and WEST-B (Washington) having the highest pass rates and the MoGEA (Missouri) and NES having the lowest pass rates. Statistical analyses suggest that candidates with comparable prior testing histories were more likely to pass the CASA than the traditional MTEL.  Candidates who took one of the CLST alternatives subsequently obtained licensure and employment at similar rates to those taking the traditional tests. MTEL-Flex Candidates participating in the MTEL-Flex pilot were about 16 percentage points more likely to pass the MTEL than qualified candidates resubmitting a traditional MTEL.  Candidates who elected to take the MTEL-Flex rather than retake the traditional MTEL were about 2-3 percentage points more likely to earn licensure and employment during the pilot.  Candidates whose scores on the MTEL qualified them for the MTEL-Flex became more likely to retake any version of the MTEL after the introduction of the alternative tests. Preparation Program Attestation Participants in the program attestation pilot were about 29 percentage points more likely to ultimately fulfill the testing requirement than other candidates in their programs taking the traditional MTEL.  Participants in the attestation pilot have made more progress toward licensure and employment than other candidates in their programs. We estimate that attestation has increased participants’ program completion rates by about 13 percentage points and employment by about 8 percentage points. |

### Licensure and Employment Data

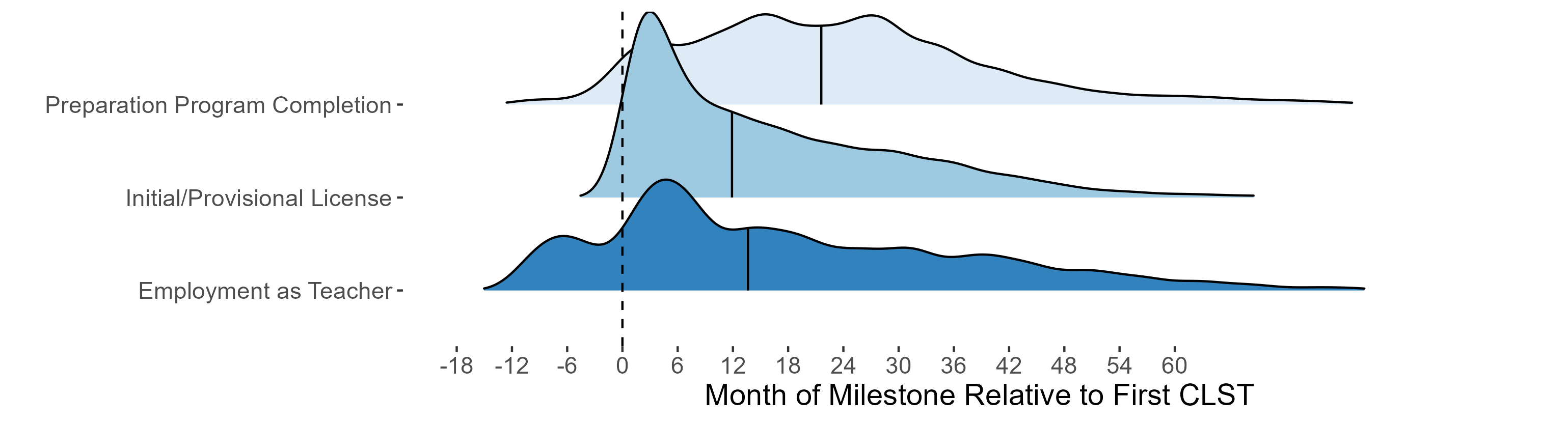
This report incorporates licensure data through September 2024 and employment data through the 2022-23 school year. We construct four primary outcomes described in Exhibit 16. These outcomes include: completion of an in-state preparation program; obtaining either an initial or provisional license to teach in Massachusetts; and employment in teaching or non-teaching roles in Massachusetts public schools.

Exhibit 16. Licensure and Employment Outcomes

|  |  |
| --- | --- |
| Outcome | Definition |
| Program Completion | Preparation program reports completion of an educator preparation program leading to initial licensure after the testing date. |
| Initial/Provisional Licensure | Candidate progresses to an initial or provisional license after the testing date. |
| Employment, Teaching Role | Candidate obtains teaching position in Massachusetts public schools in a school year following year of testing date. |
| Employment, Non-Teaching Role | Candidate obtains teaching *or* non-teaching position in Massachusetts public schools in a school year following year of testing date. |

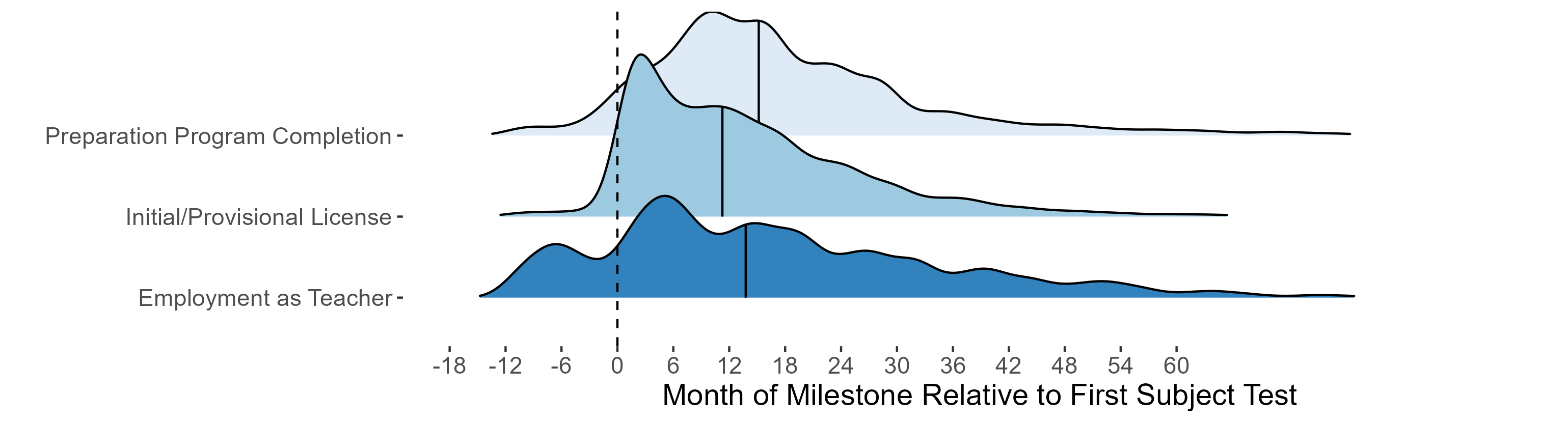
To provide some sense of the expected career progression of participants in the MTEL alternatives pilot period, we plot the timeline of key milestones relative to a candidate’s first attempt at a traditional MTEL Communications and Literacy Skills Test (Exhibit 17) or a subject test (Exhibit 18) during the pilot. In each figure, we plot the distribution of the time between (a) testing and program completion, (b) testing and initial licensure, and (c) testing and employment.

Exhibit 17. Timeline of Milestones Relative to First Communication and Literacy Skills Test Date



Candidates often take the CLST before entering a program. The median candidate first takes the CLST about 21 months prior to completing a preparation program. The median time between taking the CLST and both employment and licensure is about 1 year. Many people earn a license and become employed very shortly after taking the CLST; these are typically candidates entering the profession on a provisional license, which does not require completion of a preparation program prior to entry.

Exhibit 18. Timeline of Milestones Relative to First Subject Test Date



Candidates typically take the MTEL subject tests later in their preparation, often before their practicum. Candidates typically take the tests about 14 months before they complete a program, 11 months before they earn a license, and 10 months before becoming employed.

We plot these timelines to emphasize that many pilot participants are still progressing through the licensure system. This is particularly the case given the large number of participants in all pilot options during the 2023-24 academic year. The licensure and employment outcomes we report may continue to evolve as candidates reach these expected milestones.

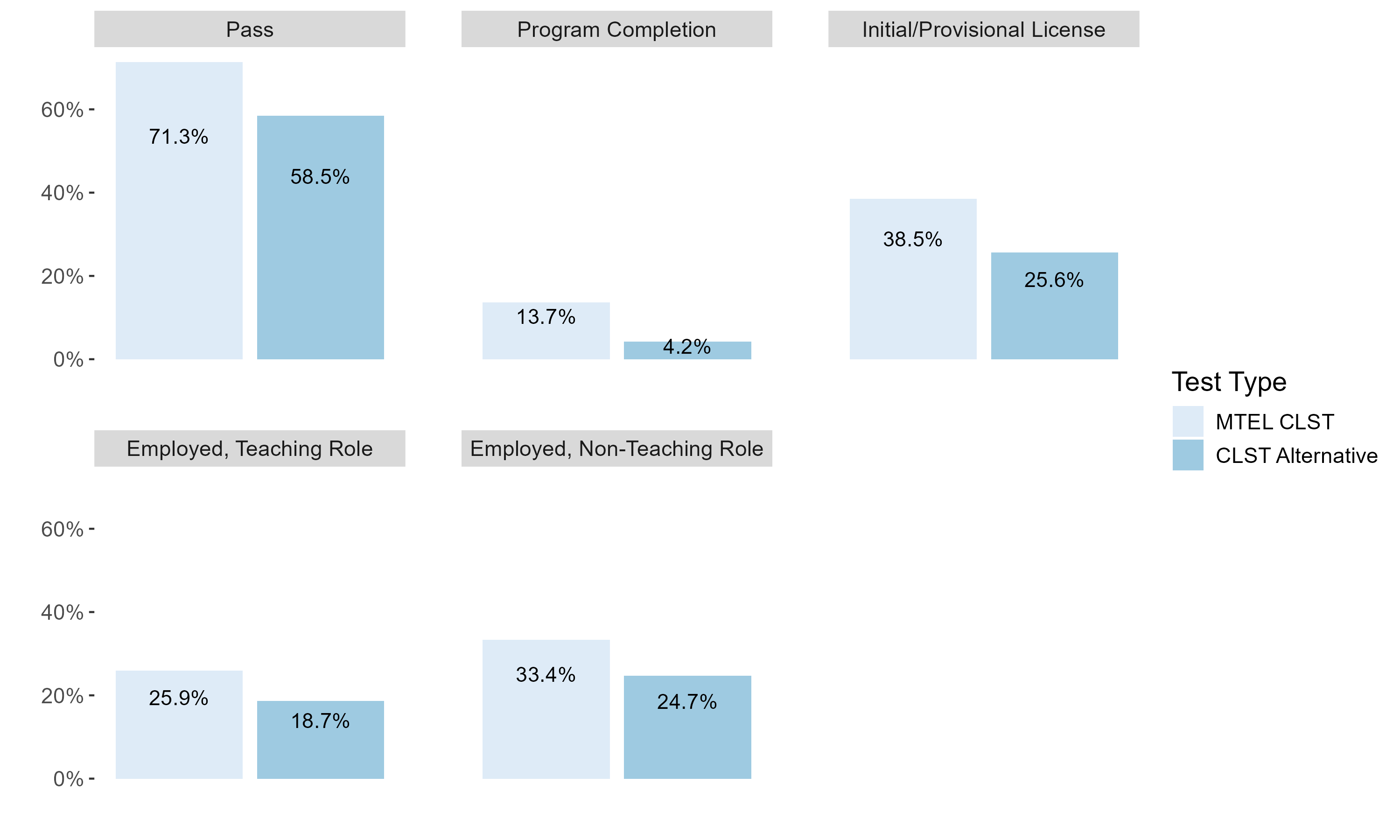
### Empirical Methods

We assess candidate licensure and employment outcomes in two ways. We first directly compare licensure and employment progress of participants taking an alternative assessment to candidates taking the traditional MTEL. However, because candidates taking alternative assessments differed in several dimensions from those taking the traditional MTEL, we also use statistical methods to adjust differences in career progression for differences in candidate background between the two groups. In particular, we regress each of the licensure and employment outcomes on several characteristics measured prior to taking the tests, including prior test-taking history.[[5]](#footnote-6) These adjustments provide evidence on how the outcomes for pilot participants differed from observationally similar candidates taking the traditional MTEL.

### Communications and Literacy Skills Alternatives

In the 2023 Annual Report, we provided preliminary evidence on career progression for candidates taking the CLST alternatives. We update this analysis with additional cohorts of test takers and additional data on program completion, licensure, and employment. In Exhibit 19, we show the percentage of traditional MTEL test-takers and those taking the CLST alternatives who reached key milestones. The overall pass rate on the CLST alternatives was about 13 percentage points lower than on the traditional MTEL (59% compared to 71%).[[6]](#footnote-7) Similarly, those taking the CLST alternatives were less likely to reach other milestones.

Exhibit 19. Candidate Outcomes on the Traditional CLST MTEL and CLST Alternatives



*Notes:* Average employment and licensure outcomes for traditional (light blue) and alternative (dark blue) MTEL test-takers. Sample includes all submissions on traditional or alternative CLSTs between October 20, 2020, and June 30, 2024.

Most of the differences in Exhibit 19, however, are explained by differences in the pool of candidates taking each assessment type. As shown in Exhibit 10, candidates with prior unsuccessful attempts on the traditional CLST MTEL were more likely to take one of the alternatives. Candidates who had previously failed an MTEL were generally less likely to meet key licensure milestones at baseline.

It is important to adjust for prior background to avoid attributing these patterns to the alternative assessments themselves. When we adjust for prior testing history and other background variables in Exhibit 20, the CLST alternatives had pass rates that were about 4 percentage points higher than the traditional MTEL CLST, and the differences in licensure and employment outcomes mostly disappear. However, one important caveat is that these findings depend on the available administrative data to adjust for differences in the kinds of teachers who opted into the two testing types (traditional MTEL CLST and CLST alternatives). In particular, the alternatives attracted candidates attending out-of-state programs or who may have been considering licensure in other states, and Massachusetts collects limited data about candidates’ state of residence. Given that we are missing information about geographical preferences, the estimates may understate the true effect of the pilot assessments on enrollment and completion for alternative test-takers to the extent that there were unobserved differences in candidates’ background or propensity to train in in-state institutions.

Exhibit 20. Differences in Workforce Outcomes Adjusted for Candidate Background

|  | Pass Test | Full License | Program Completion | Employed | Employed as Teacher |
| --- | --- | --- | --- | --- | --- |
| Panel A. All CLST Alternatives | | | | | |
| CLST Alternative | 0.040\*\* | -0.019 | -0.009 | -0.009 | 0.004 |
|  | (0.016) | (0.014) | (0.008) | (0.013) | (0.012) |
| Num.Obs. | 13969 | 13969 | 13969 | 13969 | 13969 |
| Panel B. Individual CLST Alternatives | | | | | |
| NES | -0.132\*\*\* | -0.027 | -0.014 | 0.025 | 0.048 |
|  | (0.030) | (0.035) | (0.019) | (0.032) | (0.031) |
| CASA | 0.152\*\*\* | 0.002 | -0.007 | 0.003 | -0.012 |
|  | (0.026) | (0.022) | (0.012) | (0.021) | (0.019) |
| MoGEA | 0.008 | -0.020 | -0.007 | -0.027 | -0.001 |
|  | (0.026) | (0.023) | (0.011) | (0.020) | (0.018) |
| WEST-B | 0.125 | 0.001 | 0.047 | -0.107\* | -0.052 |
|  | (0.087) | (0.076) | (0.043) | (0.059) | (0.049) |
| Praxis | 0.070 | -0.090\*\* | -0.032 | -0.038 | 0.013 |
|  | (0.052) | (0.042) | (0.023) | (0.042) | (0.042) |
| Num.Obs. | 13,969 | 13,969 | 13,969 | 13,969 | 13,969 |
| *Notes:* Coefficients from regressions of licensure and employment outcomes on CLST choice. Sample includes CLST subsmissions during the 2020-21 and 2021-22 school years (after October 20, 2020). In addition to the included characteristics shown in in Exhibit 12, the regressions also include a quadratic polynomial of the number of days since the beginning of the pilot. All controls measured at the testing date. Standard errors clustered by teacher candidate. \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 | | | | | |

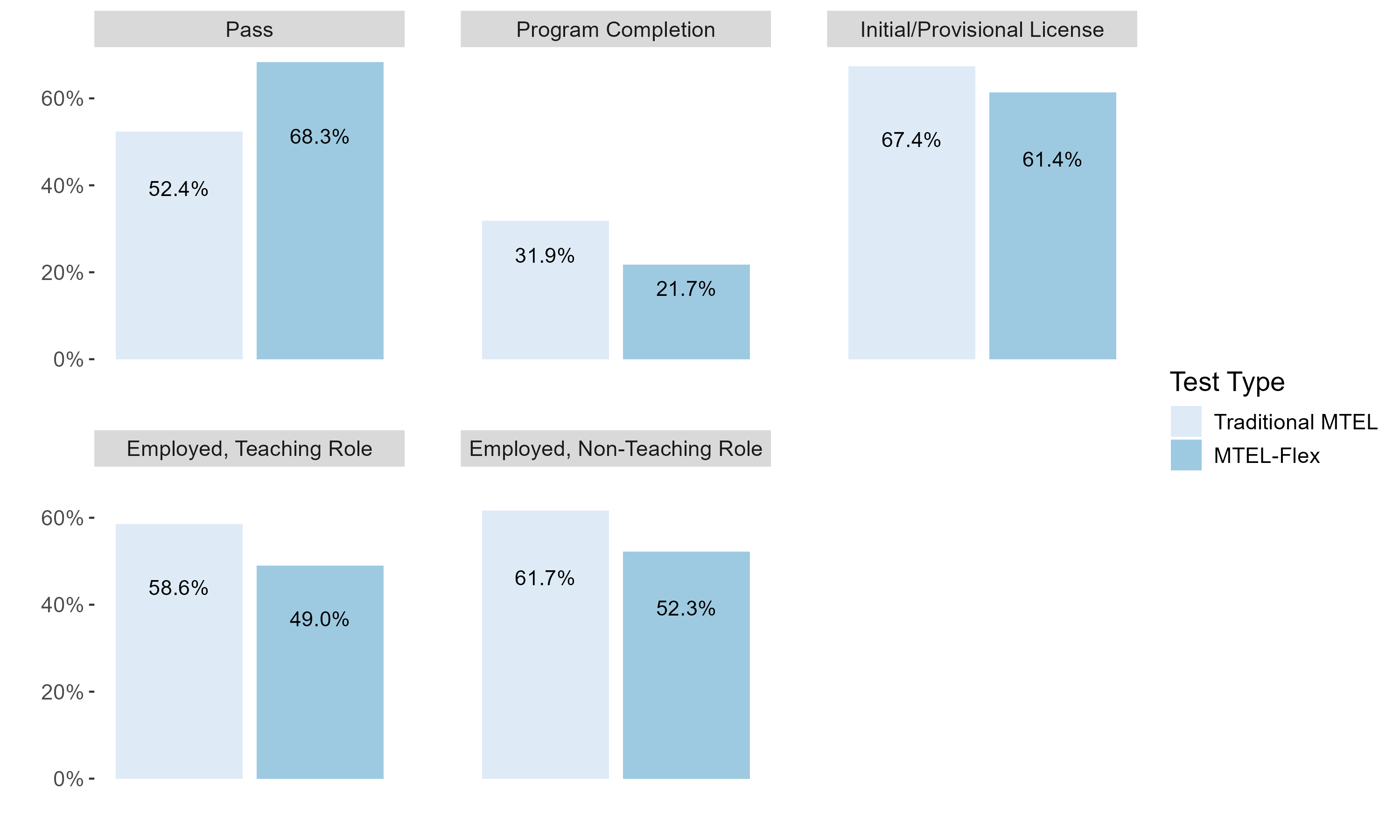
In Panel B of Exhibit 20, we estimate differences in testing and workforce outcomes by the test series. Each coefficient describes differences in these outcomes between candidates taking a reading or writing subtest in one of the indicated series and those taking the traditional CLST. We do observe some differences in outcomes across the testing series. The NES appears to have been more difficult than the traditional MTEL: the pass rate on the NES was about 13 percentage points lower than the MTEL conditional on prior testing history. The CASA and WEST-B had higher pass rates by 15 and 13 percentage points, respectively, although the result for the WEST-B is not statistically significant as relatively few candidates had taken this test. We generally do not find significant results on other outcomes, although we cannot rule out meaningful effects on licensure or employment outcomes.

The CASA was both the most popular testing series among the piloted alternatives (Exhibit 3) and the assessment with the highest passing rate (Exhibit 20). The relative difficulty of the assessment is one explanation for its popularity, particularly because the test is only accepted for licensure in Massachusetts and Indiana. However, the candidate surveys suggest that there may also be other explanations. In particular, the MoGEA, which appears to be comparable in difficulty to the MTEL CLST, was the second most taken test series. Like the CASA, however, this test is less costly than the MTEL and is offered in an online-only format. This suggests that cost and accessibility may also play important roles in candidates’ choice of test. Regardless of the reasons behind choice of test, there do appear to be meaningful differences in the difficulty of the assessments.

### MTEL-Flex

As shown in Exhibit 21, candidates who took the MTEL-Flex were about 16 percentage points more likely to have passed the subject test than candidates retaking the MTEL. In Exhibit B-2, we show that pass rates on MTEL-Flex are significantly higher relative to the traditional MTEL for Asian and Pacific Islander, Hispanic, and Black candidates. The regression-adjusted difference in Exhibit 22 that controls for candidate background is similar, suggesting that taking the MTEL-Flex did increase the likelihood of passing relative to the traditional MTEL. This difference has increased over time; we estimated that the MTEL-Flex pass rate was about 6-10 percentage points higher than the traditional MTEL on the 2022 and 2023 Annual Reports. We return to changes in the MTEL-Flex pass rate below.

Exhibit 21. Candidate Outcomes on the MTEL-Flex



*Notes:* Average licensure and employment outcomes for candidates taking traditional MTEL (light blue) and MTEL-Flex (dark blue).

Exhibit 22. Differences in Workforce Outcomes Adjusted for Candidate Background

|  | Pass Test | Full License | Program Completion | Employed | Employed as Teacher |
| --- | --- | --- | --- | --- | --- |
| MTEL-Flex | 0.161\*\*\* | 0.034\*\* | 0.012 | 0.021 | 0.022\* |
|  | (0.015) | (0.016) | (0.010) | (0.013) | (0.013) |
| Num.Obs. | 5677 | 5677 | 5677 | 5677 | 5677 |
| *Notes:* Coefficients from regressions of teacher candidate outcomes on MTEL-Flex test choice. Sample includes all candidates whose scores on a prior MTEL submission qualify them to take the MTEL-Flex and who subsequently retook the test. Program enrollment, program completion, and initial/provisional license are all measured after the testing date. Employment outcomes are measured during academic years following the year a candidate qualified for the test. In addition to the included characteristics, the regressions also include a quadratic polynomial of the number of days since the beginning of the pilot. All characteristics measured at the testing date. Standard errors clustered by teacher candidate. \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 | | | | | |

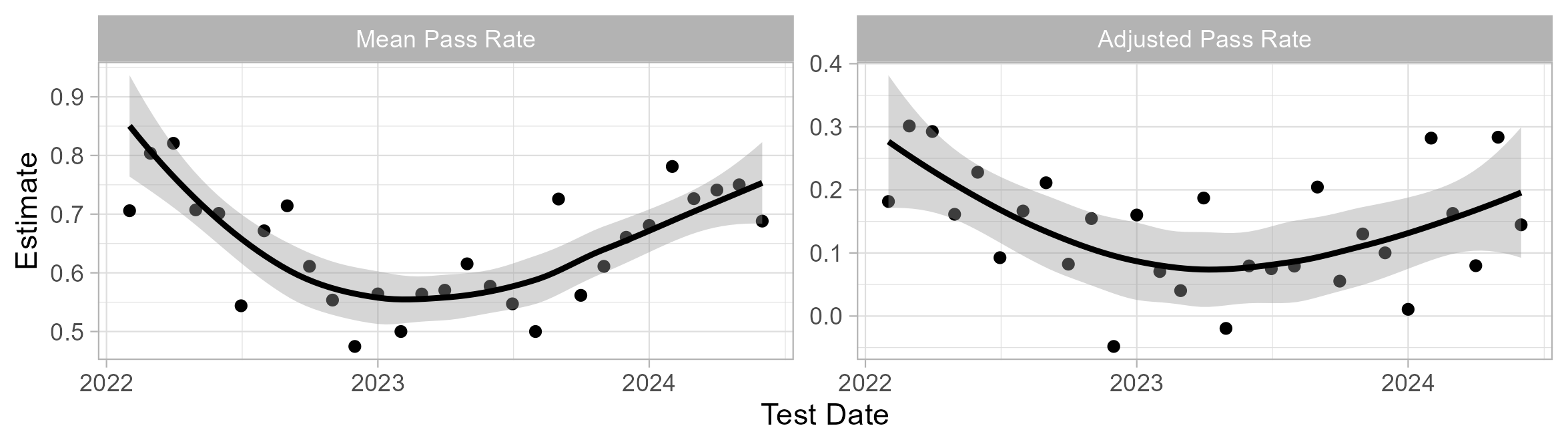
Consistent with the higher pass rates for the MTEL-Flex participants, we also observe higher licensure and employment outcomes. MTEL-Flex participants were about 3 percentage points more likely to have earned an initial or provisional license. They were about 2 percentage points more likely to be working as a teacher (significant at the 10% level). MTEL-Flex participants also had similar rates of program completion following the test submission, although we note that they were disproportionately likely to already be working as educators on emergency licenses.

#### Variation in MTEL-Flex Pass Rates During the Pilot

Pass rates on the MTEL-Flex have increased relative to the traditional MTEL over the course of the pilot. In the interim reports, we found that candidates who took the MTEL-Flex had pass rates about 6-10 percentage points higher than those taking the traditional MTEL after adjusting for differences in candidate background. The difference in pass rates has since increased to about 16 percentage points. We plot variation in the pass rates on the MTEL-Flex during the pilot in Exhibit 23. The left-hand plot displays average pass rates by testing month. The right-hand plot shows estimates of the difference in pass rates between MTEL-Flex and the traditional MTEL after adjusting for differences in candidate background. Positive values indicate that candidates with similar backgrounds were more likely to pass the MTEL-Flex.

In both cases, we observe relatively high pass rates among the earliest and most recent participants. During the first several months of the MTEL-Flex, pass rates were in the 70-80% range and about 25 percentage points higher than comparable candidates taking the traditional MTEL. Pass rates declined to about 55-60% by the end of 2022 and increased back above 70% by mid-2024.

Exhibit 23. MTEL-Flex Pass Rates by Test Month



*Notes:* Average pass rate and adjusted pass rate on MTEL-Flex by test month. The adjusted pass rate is an estimate of the difference between pass rates on the MTEL-Flex and traditional MTEL adjusted for prior performance on the MTEL. The adjusted pass rate is derived from a regression of test results on an interaction between MTEL-Flex and testing month using the sample of MTEL-Flex-eligible resubmissions.

#### MTEL-Flex Qualification and Test Resubmissions

The analyses above consider differences in pass rates, but the MTEL-Flex may have increased licensure rates through other channels as well. In particular, the availability of the MTEL-Flex may have encouraged candidates who would not have otherwise retaken the MTEL. About 30% of teacher candidates who barely fail the MTEL do not make another attempt on the test (Cowan et al., 2024). Retest rates are especially low for teacher candidates of color. The MTEL-Flex provided another option for these candidates and may have affected whether they make another attempt at the test. In this section, we explore how MTEL-Flex eligibility affected candidates’ decision to reattempt the MTEL.

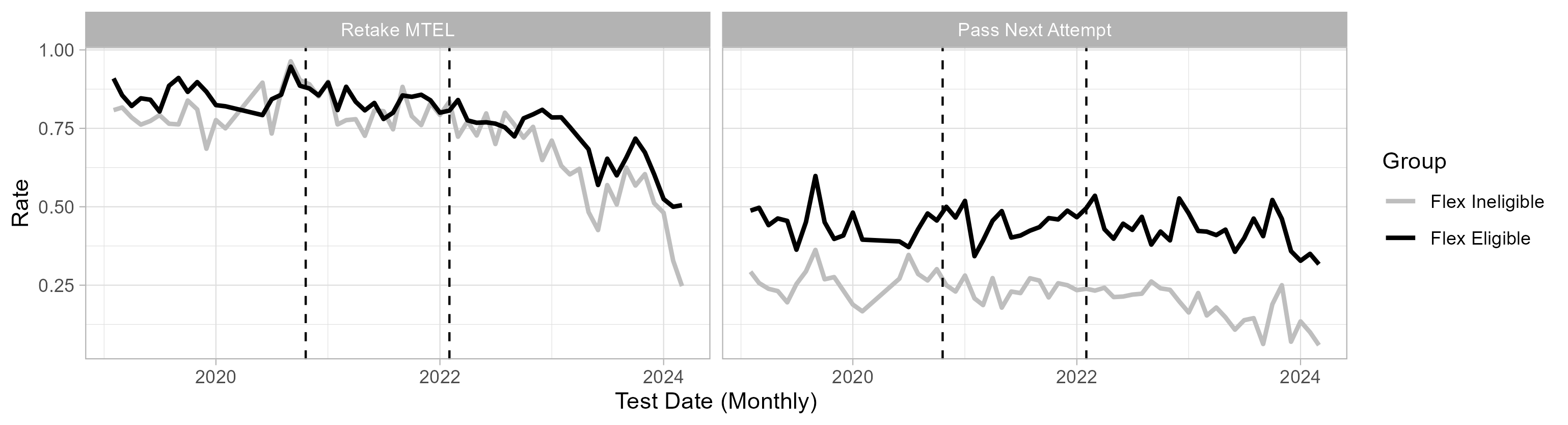
In order to understand whether qualification for the MTEL-Flex affected candidates’ testing decisions, we construct a comparison group of teacher candidates who were not eligible to take the MTEL-Flex. We then compare retesting decisions for this control group to a treated group of candidates eligible for MTEL-Flex. As a control group, we use teacher candidates who scored just below the MTEL-Flex eligibility threshold. Hence, our treated group consists of candidates who received scores of 0-1 SEM below the MTEL passing threshold, and our control group consists of candidates who scored 1-2 SEM below. We then compare changes in retesting outcomes for these two groups before and after the introduction of the MTEL-Flex.

We construct two outcomes for each candidate: (1) an indicator for whether the candidate retook the same MTEL; and (2) an indicator for whether the candidate retook the same MTEL *and passed* *on the next attempt*. We focus on the tests that became available in early 2022 (English, History, Spanish, English as a Second Language, General Curriculum, and Foundations of Reading) because later tests either coincide with a change in the traditional MTEL, have few post-eligibility time periods, or both.

We summarize these outcomes for each ethnoracial identification group in Appendix Exhibit B-3. Although about 77% of candidates in this sample retook either the MTEL or MTEL-Flex, the pass rates differed significantly across groups: retesting rates were 78% for white candidates and 79% for Asian American candidates; by contrast, 71% of Hispanic candidates and 66% of Black candidates retook the test. Other research has shown that differences in these “walk-away” rates contribute meaningfully to disparities in eventual passing status (Cowan et al., 2023).

We next plot the average outcomes for all candidates by month in Exhibit 24 for MTEL-Flex eligible scores (black) and the ineligible controls (gray). The dashed vertical lines indicate the beginning of the MTEL alternatives pilot and the first month of MTEL-Flex submissions. Resubmission rates were similar for eligible and ineligible scores prior to the introduction of the MTEL-Flex in February 2022. Starting in the middle of 2022, the retake rates began to diverge between eligible and ineligible scores. Retake rates for MTEL-Flex-eligible scores remained higher throughout the remainder of the pilot. Trends in passing outcomes are less clear; eligible scores were always more likely to retake and pass on the next attempt than ineligible scores.

Exhibit 24. Candidate Outcomes on the MTEL-Flex



*Notes:* Proportions of candidates who retake the MTEL (left) or retake and pass the next attempt (right) by the month of the initial MTEL attempt. The MTEL-Flex Eligible group includes those scoring between 0 and 1 SEM below the minimum qualifying score; the MTEL-Flex Ineligible group includes those scoring between 1 and 2 SEM below the minimum qualifying score. The dashed lines indicate the months of the start of the pilot and the introduction of the MTEL-Flex.

The trends in Exhibit 24 provide suggestive evidence that qualification for MTEL-Flex increased resubmission rates. We explore this more formally using a difference-in-differences design that compares changes in outcomes for eligible and ineligible scores throughout the pilot period. Let *i* denote an MTEL submission, *k* denote a scaled score on the MTEL and *t* denote the month of the test. Using observations between 0 and 2 SEM below the minimum qualifying score, we estimate

(1)

where indicates that the score qualified the candidate for the MTEL-Flex and indicates that the test was taken in a month where the MTEL-Flex was available. The coefficient describes how outcomes changed for candidates receiving qualifying scores after the introduction of the MTEL-Flex compared to changes for candidates whose scores do not quite qualify them for the MTEL-Flex.[[7]](#footnote-8)

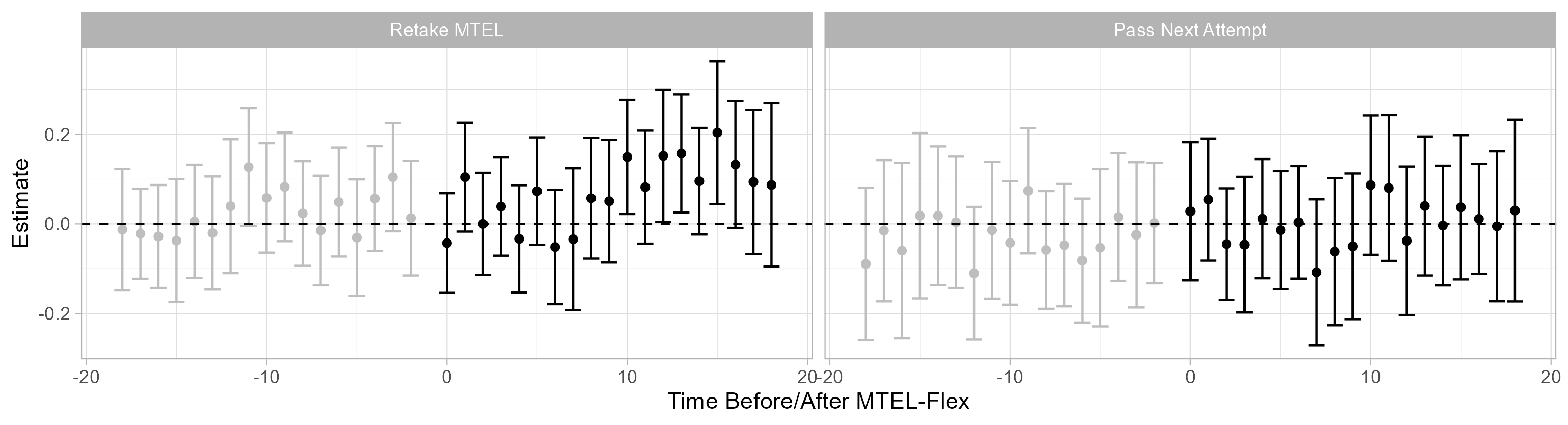
We show estimated effects of MTEL qualification on retesting in Exhibit 25. Overall, we estimate that qualification for the MTEL-Flex increased retesting rates by about 3 percentage points (significant at the 10% level). We do not find any statistically significant effects on passing the MTEL, although the point estimates are similar. The trends shown in Exhibit 24 indicate that retest rates diverged between the MTEL-Flex-Eligible and MTEL-Flex-Ineligible groups some months after the introduction of the MTEL-Flex. In the second two columns, we let the effect of MTEL-Flex qualification differ with the time since the introduction of the test. We find little effect of MTEL-Flex qualification during the first year of the MTEL-Flex. After 12 months, we estimate an effect of about 10 percentage points on retesting.

Exhibit 25. MTEL-Flex Qualification and Retesting

|  | Retake | Pass Next | Retake | Pass Next |
| --- | --- | --- | --- | --- |
| Flex Eligible | 0.029\* | 0.032 |  |  |
|  | (0.017) | (0.023) |  |  |
| Flex Eligible (0-11 months) |  |  | -0.014 | 0.029 |
|  |  |  | (0.016) | (0.023) |
| Flex Eligible (12+ months) |  |  | 0.096\*\*\* | 0.037 |
|  |  |  | (0.028) | (0.030) |
| Num.Obs. | 15483 | 15483 | 15483 | 15483 |
| *Notes:* Difference-in-differences estimates of the effects of MTEL-Flex qualification on retest and passing outcomes. Standard errors clustered by MTEL scale score. \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 | | | | |

We plot the estimated effects by month relative to the introduction of the MTEL-Flex in Exhibit 26. The effects for negative periods (gray) indicate that there is little evidence that retesting behavior in the MTEL-Flex-Eligible and MTEL-Flex-Ineligible groups was diverging prior to the introduction of the MTEL-Flex. Consistent with Exhibits 24 and 25, we see an increase in retesting among the eligible group starting about 8 months after the introduction of MTEL-Flex.

Exhibit 26. Dynamic Effects of MTEL-Flex Qualification on Retesting



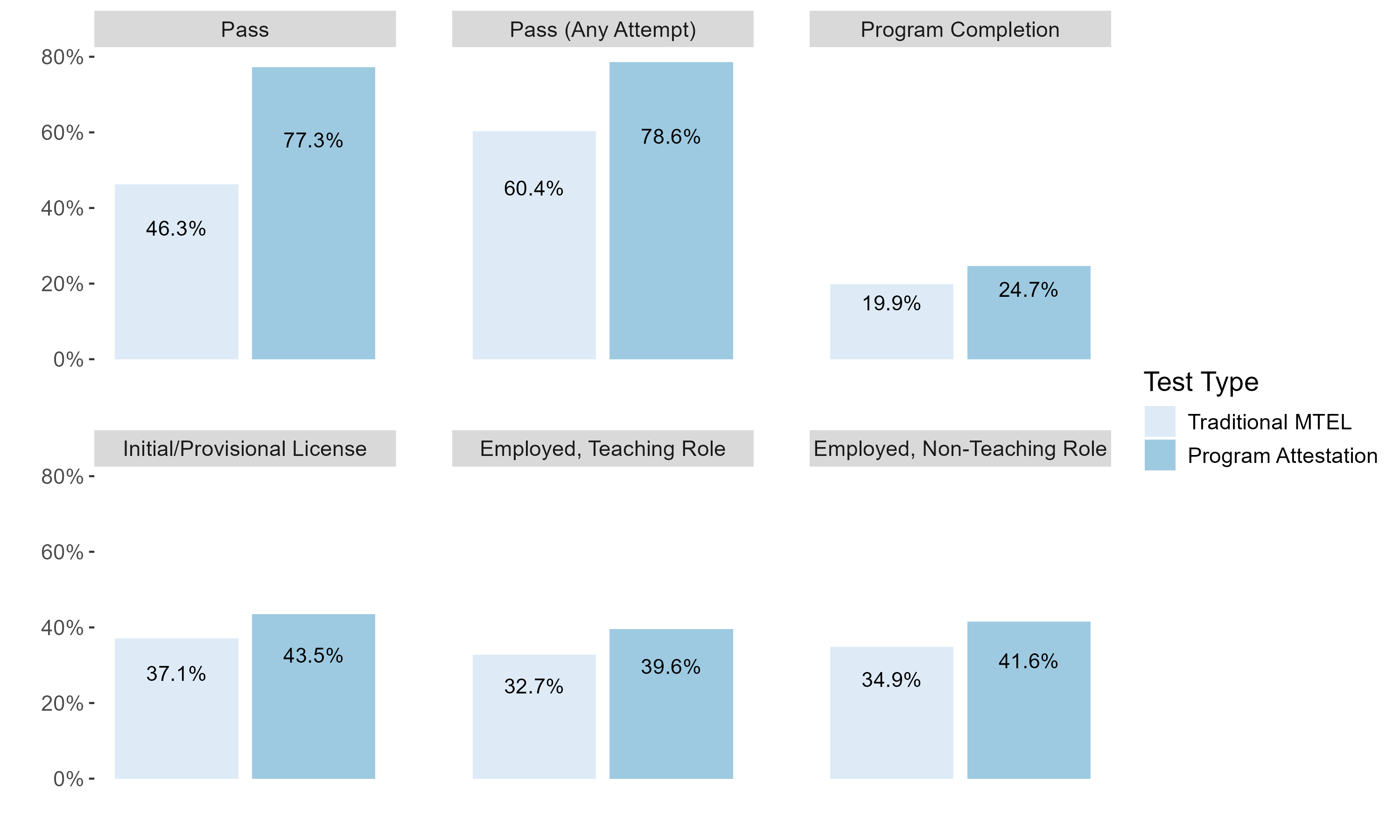
*Notes:* Estimated effects of MTEL qualification on retaking the MTEL (left) or retaking and passing the next attempt (right) by the month relative to the introduction of the MTEL-Flex. The MTEL-Flex Eligible group includes those scoring between 0 and 1 SEM below the minimum qualifying score; the MTEL-Flex Ineligible group includes those scoring between 1 and 2 SEM below the minimum qualifying score. The monthly effects are estimated by replacing the indicator in Eq. (1) with an indicator for each month relative to the introduction of the MTEL-Flex. 95% confidence intervals constructed using standard errors clustered by MTEL scale score.

At first glance, the null findings on whether MTEL-Flex-Eligible candidates passed their next attempt may seem inconsistent with the findings on overall pass rates in Exhibit 22. To put these findings in perspective, we estimate a back-of-the-envelope effect on passing based on the findings on pass rates (Exhibit 22) and retakes (Exhibit 25). Assuming pass rates on the MTEL-Flex of 68%, a relative difference in pass rates between the MTEL-Flex and traditional MTEL of 16 percentage points, an increase in retakes of 3 percentage points, and that 24% of eligible candidates take the MTEL-Flex, we would expect an effect on retaking the MTEL and passing the next attemptof about 5 percentage points. Although the difference-in-differences estimate (3 percentage points) is not significant, the point estimate is close to what we would expect given the other results in this report. In other words, the insignificant finding may be explained by small sample sizes rather than there being no effect of MTEL-Flex qualification on pass rates.

### Program Attestation

In Exhibit 27, we plot candidate outcomes for participants in the program attestation pilot. We compare them to other candidates that enrolled in preparation programs offered by the same provider and who submitted scores on the same traditional MTELs offered through the attestation option. Pass rates on the program attestation option (77%) declined as the programs have expanded, but they remained significantly higher than the traditional MTEL throughout the pilot (46%). However, the attestation option allows candidates to resubmit materials after receiving feedback from program faculty. The pass rates are therefore not exactly comparable to one-time pass rates on the traditional MTEL. In the next plot, we therefore show results for whether a candidate passed *any* attempt. The difference in pass rates declines from 31 percentage points to 18 percentage points, suggesting that program attestation both improved overall pass rates and reduced the number of attempts candidates make on the MTEL. That said, program administrators reported that the attestation alternative was more intensive than the traditional MTEL and took up to an entire academic term to complete (see Appendix A). Hence, the reduction in test attempts may be offset by additional preparation for the attestation process. In the remaining plots, we show that rates of program completion, licensure, and employment were all somewhat higher for attestation participants.

Exhibit 27. Percentage of Program Attestation Candidates Attaining Licensure/Employment Milestones



As shown in the section on candidate background, the attestation participants differed on prior testing and licensure outcomes from candidates taking the traditional MTEL. We adjust for these differences in Exhibit 28 by matching candidates in the attestation pilot to candidates who took the traditional MTEL in their preparation programs.[[8]](#footnote-9) We then compare candidate outcomes for pilot participants to the matched sample. The results are shown in Exhibit 28.

Exhibit 28. Differences in Workforce Outcomes Adjusted for Candidate Background

|  | Pass Test | Pass Test  (Any Attempt) | Full License | Program Completion | Employed | Employed as Teacher |
| --- | --- | --- | --- | --- | --- | --- |
| Attestation | 0.454\*\*\* | 0.290\*\*\* | 0.076 | 0.129\*\*\* | 0.057 | 0.078\* |
|  | (0.055) | (0.054) | (0.055) | (0.046) | (0.047) | (0.046) |
| Num.Obs. | 234 | 234 | 234 | 234 | 234 | 234 |
| \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 | | | | | | |

Because participants in the attestation option were more likely to have previously taken and failed an MTEL, the adjusted pass rates are larger than their counterparts in Exhibit 27. We estimate that participating in the attestation option increased pass rates by 29 percentage points, program completion by about 13 percentage points, and employment by about 8 percentage points (significant at the 10% level).

## Teacher Effectiveness

|  |
| --- |
| Key findings Communication and Literacy Skills Tests Alternatives Teachers who passed a CLST alternative and who entered the workforce were similarly effective as candidates who took the traditional MTEL based on performance ratings and overall contributions to student achievement on MCAS. They may have been less effective in math instruction and at improving a battery of non-tested outcomes than teachers passing the CLST MTEL, but these differences were not statistically significant at conventional levels. MTEL-Flex Teachers who passed an MTEL-Flex and entered the workforce made similar contributions to student growth on MCAS, non-tested outcomes, and school climate as teachers passing the traditional MTEL subject tests.  Teachers who passed the MTEL-Flex were less likely to earn exemplary ratings (the highest performance category in Massachusetts) than candidates passing the traditional MTEL subject tests, but they were no more likely to earn low performance evaluations (unsatisfactory or needs improvement). Preparation Program Attestation We do not find that teachers whose preparation providers attested to their subject matter knowledge were less effective than other candidates from their programs, but the small number of participants in this option means that we cannot rule out significant differences in performance evaluations. |

In this section, we use data on teachers who were employed in the 2020-21 through 2023-24 school years and their students to compare the effectiveness of teachers who passed an alternative assessment to those passing the traditional MTEL. We collect several student outcomes and estimate value-added models that aim to estimate the effects of being assigned to teachers with different attributes on achievement and other outcomes (Bacher-Hicks & Koedel, 2023). We also collect teacher evaluation ratings to examine the performance of pilot participants. In the next section, we discuss the data and empirical methods before presenting results for each of the pilot assessments.

### Data and Empirical Methods

We use four sources of data to measure teaching effectiveness in this report, which we summarize in Exhibit 29. First, we use student test scores on the Massachusetts Comprehensive Assessment System (MCAS). MCAS is administered annually in math and ELA in grades 3 – 8 and 10. Using classroom assignment data, we link students in grades 4 – 8 and 10 to their math and ELA teachers and use MCAS data to estimate teacher effects on student achievement. We supplement these data with science test scores in grades 5, 8, 9, and 10 (see Appendix B). We standardize MCAS data by grade and year so that scores are comparable across grade level and time.

We also construct a measure of students’ academic behaviors using data on absences, suspensions, on-time grade progression, and course grades.[[9]](#footnote-10) We estimate a factor model to obtain a summary measure of student behaviors that is a weighted average of the four academic outcomes; we then standardize the factor scores by grade and school year and use it as an additional outcome in the value-added models. Prior work in Massachusetts and elsewhere has found that these measures predict teachers’ effects on long-run outcomes, including high school graduation and college enrollment (Backes et al., 2024a; Jackson, 2018).

We complement these data with student surveys of academic climate. The Views of Climate and Learning (VOCAL) survey is administered annually to students in grades 4, 5, 8, and 10. The survey covers several topics related to engagement in school, social interactions with students and faculty, physical and emotional safety, and discipline. DESE creates a summary measure of students’ VOCAL responses to all survey questions. We standardize the summary measure by grade and school year and use this as an outcome in our value-added models. Prior work has shown that student surveys provide meaningful information about teacher effectiveness and classroom practice (Backes et al., 2022).

Finally, we use annual performance evaluations. Teachers and other educators receive performance evaluations under the Massachusetts Educator Evaluation Framework. The evaluations are typically conducted by school administrators and incorporate several sources of information about teacher performance. Each educator receives a summary rating on a four-point scale (unsatisfactory, needs improvement, proficient, and exemplary). We construct a summary measure of teacher performance by fitting an item response model to the overall performance rating. We use the teacher scores from these models as a measure of annual performance.[[10]](#footnote-11) Because 80-90% of teachers receive a proficient rating each year, we separately identify teachers with exceptionally low (unsatisfactory or needs improvement) or high (exemplary) ratings.

Exhibit 29. Licensure and Employment Outcomes

|  |  |  |
| --- | --- | --- |
| Outcome | Years Available | Definition |
| Math, ELA, and Science MCAS Scores | 2021-2024 | Standardized test scores on MCAS in grades 4 – 8 and 10 (5, 8, 9, and 10 for science) |
| Summary Nontest Index | 2021-2023 | Summary measure of student academic behaviors using data on attendance, suspensions, on-time grade progression, and course grades |
| Views of Climate and Learning (VOCAL) Survey | 2021-2024 | Student survey of engagement and school climate |
| Teacher Performance Evaluations | 2021-2023 | Annual performance evaluations |

We compare the effectiveness of those who passed a piloted alternative assessment to other groups of Massachusetts educators for the three student measures using value-added models estimated from data that links students to their teachers. For each outcome, we estimate the regression

. (2)

In Eq. (2), *i* denotes individual students, *j* denotes teachers, *t* indicates year, and *s* indicates school. is a set of control variables, including prior-year math and ELA test scores, prior-year nontest outcomes, student demographics, and classroom characteristics. is a variable indicating that teacher *j* passed one of the MTEL alternative assessments. is a set of other teacher characteristics. This always includes teacher experience because participants in the MTEL pilots tend to be early career teachers. Finally, is a school-by-year effect so that teachers who passed a piloted alternative assessment are compared to other teachers working in their schools. The descriptive analyses in this report show that pilot participants were more likely to enter the profession on an emergency license and have more frequently failed prior attempts on the MTEL. These teachers tend to work in lower achieving schools (Backes et al., 2024b). In addition, novice teachers are generally more likely to work in high-needs schools than more experienced teachers. Lower-achieving schools tend to have fewer resources and poorer student outcomes for reasons unrelated to the effectiveness of their teacher workforce (Mansfield, 2015). Failing to account for school assignments would attribute differences in school resources or school effectiveness to individual teachers, which would tend to artificially reduce the apparent effectiveness of pilot participants. Comparing pilot participants to others in their schools removes this potential source of bias.

We use teacher-level data to compare the performance evaluations of pilot participants and other educators in Massachusetts. For each summative performance rating outcome (), we estimate the regression

. (3)

In Eq. (3), is a vector of controls for teacher assignments, including subject and grade effects, an indicator for whether the evaluation is a formative assessment, and an indicator for whether the teacher is new to their role in the school that year.[[11]](#footnote-12) As before, is a variable indicating that teacher *j* passed one of the MTEL alternative assessments, is a set of other teacher characteristics, and is a school-by-year effect. As in Eq. (2), including school-by-year effects yields comparisons of pilot participants to other teachers in their same school. We limit comparisons to other teachers in the same school because evaluators vary in their stringency applying performance ratings to the same teacher (Cowan et al., 2022).

We assess the effectiveness of pilot participants in order to draw inferences about how the pilots might have affected the composition of the teacher workforce in Massachusetts. Identification of a policy-relevant comparison for teacher candidates participating in the MTEL pilots and passing the alternative assessments is a key challenge to this exercise. An ideal experiment would allow us to observe who would be employed in the absence of the availability of the MTEL alternatives. We could then directly compare the effectiveness of teachers who enter the profession due to the pilot assessments to other educators. Because this is infeasible, we must instead rely on comparisons with other teachers in the workforce that are arguably good proxies for this unobserved counterfactual.

We start by comparing teachers who completed licensure requirements via the alternative assessments to all other teachers in Massachusetts with comparable prior teaching experience who were observed with the same outcome measures (performance evaluation ratings, value-added effectiveness, etc.).[[12]](#footnote-13) This specification only includes controls for years of teaching experience. Teachers improve rapidly during their first few years in the workforce, and pilot participants tend to be early-career teachers (Clotfelter et al., 2007; Papay & Kraft, 2015; Rockoff, 2004). Because comparing younger pilot participants to more seasoned teachers is not informative of how each set of teachers would perform at the same experience level, we adjust for differences in teacher experience between the pilot candidates and other teachers in the state. This adjustment limits the comparison to other teachers in Massachusetts with similar levels of experience. This includes all other teachers, regardless of whether they had taken an MTEL during the pilot period or were licensed in one of the areas covered by the pilot assessments.

We provide this comparison for descriptive purposes, but we note that these results will be sensitive both to implementation decisions (e.g., which tests were included in the pilot) and broader contextual factors (e.g., which teacher candidates were actively taking licensure tests). The subset of tests included in the MTEL-Flex and program attestation options are not random, and teacher effectiveness may vary across licensure areas. For instance, the English language learners assessment was overrepresented in both the MTEL-Flex and program attestation pilots.

Perhaps more importantly, the MTEL alternatives pilot also coincided with emergency licensure, which was a significant intervention in licensure policy in Massachusetts (Backes et al., 2024b). The emergency licensure policy temporarily exempted teachers from testing requirements. To continue teaching on a license after their emergency license expired, teachers on emergency licenses needed to pass MTEL. This requirement changed the composition of the licensure testing pool at the same time as the new assessments were being piloted. Other research in Massachusetts has shown that recent cohorts of emergency licensed teachers were less effective on average than other teachers with comparable teaching experience (Backes et al., 2024b). As a consequence, the pool of teacher candidates taking MTELs during the pilot period was less effective on average than other teachers with similar levels of experience.

Our preferred baseline specification therefore adds a control for whether the teacher passed a traditional or alternative MTEL during the pilot to . For the CLST alternatives, this variable includes teachers who passed either the traditional MTEL CLST or an alternative assessment during the pilot period. For MTEL-Flex, this indicates that the teacher passed one of the traditional MTEL subject tests included in the pilot or the MTEL-Flex during the pilot period. For program attestation, this indicates that the teacher passed one of the MTEL subject tests included in their program’s attestation pilot and submitted their scores to the indicated teacher preparation program. In all cases, these models compare teachers who passed the pilot assessments to those who passed the traditional MTEL. This approach mitigates the issues discussed above. Limiting the sample to those taking similar MTELs during the pilot ensures that the comparison teachers better reflect the pool of candidates who might have participated in the pilots. By including an indicator for whether a teacher passed an MTEL during the pilot period, we also ensure that comparisons between pilot participants and other educators is limited to those who might have obtained licensure via the alternative assessments.

The main limitation of this approach is that not all pilot participants would have been excluded from the profession in the absence of the pilot. If the alternative assessments were not offered, many of the participants would have passed the traditional MTEL instead. Thus, it is instructive to consider two groups of pilot participants based on their hypothetical performance on the traditional MTEL. The *always-employed* group would have passed the traditional MTEL and obtained a teaching position if the pilot were not offered. The pilot assessments might benefit this group by reducing testing costs, but their availability does not affect these teachers’ ultimate employment status. The *newly-employed* group consists of teachers who would not have passed the traditional MTEL in the absence of the pilot and only obtain teaching positions because they passed an alternative assessment. These teachers might have been unable to pass the traditional MTEL or might have dropped out rather than retake a failed assessment. The analyses of workforce outcomes indicate that both the MTEL-Flex and program attestation options increased licensure and employment rates, which suggests that both pilots had some teachers in the newly-employed group. However, these analyses also indicate that most of the teachers passing the pilot assessments would have been in the always-employed group. This distinction is important because the pilots only affected the composition of the workforce by moving the newly-employed into teaching positions.

The prior specifications include both always- and newly-employed teachers when comparing pilot participants to other educators. Although it is not possible to identify which teachers earn licensure due to the availability of the pilot assessments, we can assess whether the newly-employed group differed in effectiveness from those who would have entered the profession regardless of the pilot. To do so, we compare teachers passing the alternative assessments to observationally similar teachers who passed the traditional MTEL. In practice, we add controls for prior MTEL performance to to adjust for compositional differences between pilot participants and traditional MTEL test-takers. Assuming we successfully adjust for selection into test type, any remaining differences in teacher effectiveness between those passing the traditional MTEL and those passing the alternatives must be due either to (a) effects of the alternatives on the performance of individual teachers or (b) effects on the composition of who became a teacher.[[13]](#footnote-14) We assume that the effects of participating in the pilots on teacher performance is minimal. We believe this is a reasonable assumption, at least for the CLST alternatives and MTEL-Flex. There is little reason to believe that taking a different standardized communications and literacy skills assessment would lead to significant improvements in teaching practice. The MTEL-Flex does include some additional preparation to complete the assessment, but teachers reported spending only about 10 hours on the MTEL-Flex and the average effects of modest professional development activities on student outcomes is quite small. This assumption may be less reasonable for the attestation option, where the time investment is more substantial.

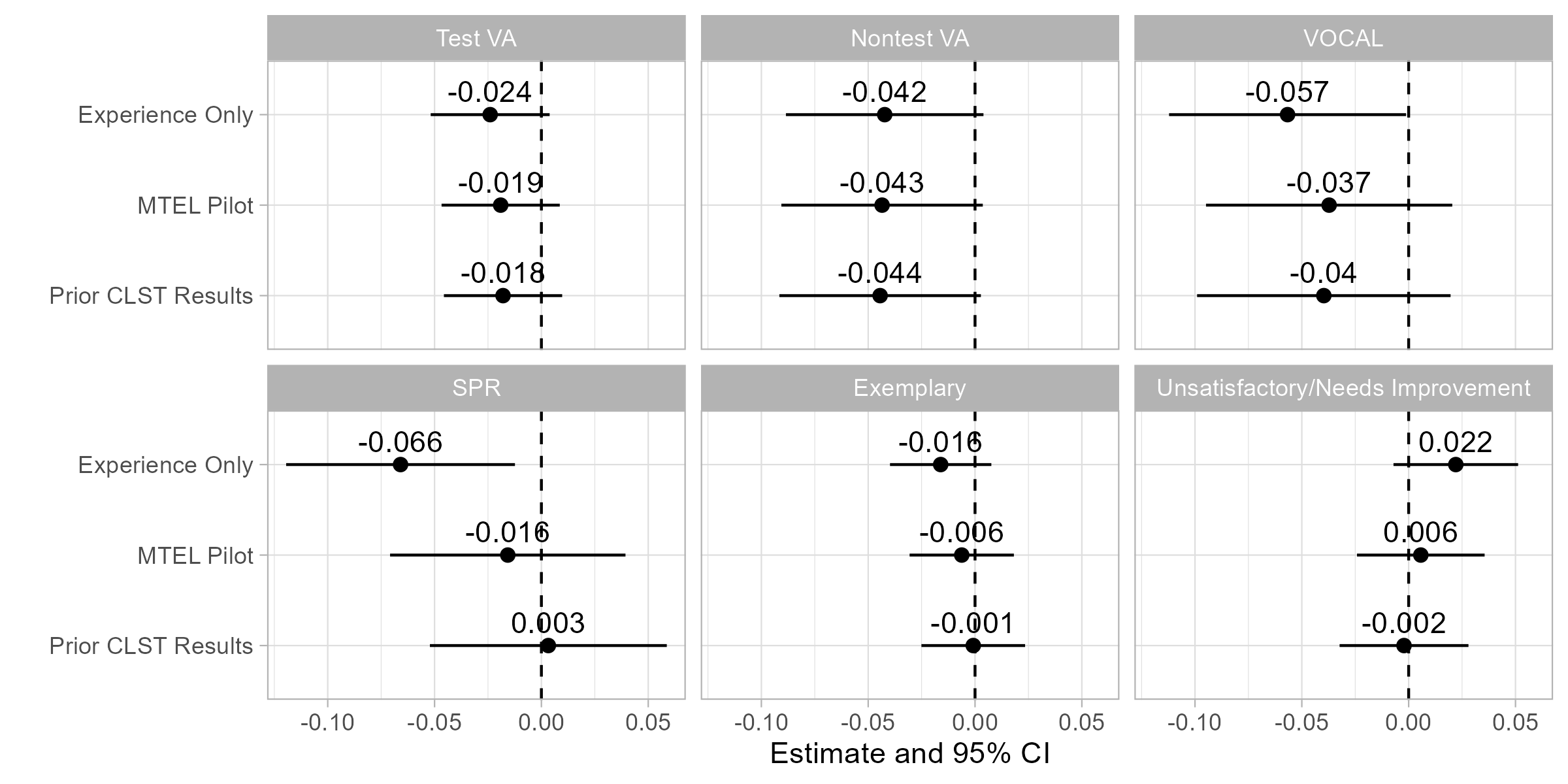
If we rule out the possibility that the alternative assessments improve teaching practice, then the second explanation for any observed differences in educator effectiveness is that the always-employed group differs from the newly-employed group. In other words, this would suggest that teachers who entered the workforce only due to the availability of the alternative assessments are differently effective than those who would have obtained licensure in their absence. We therefore interpret these specifications as a test that the two hypothetical groups of teachers are differently effective. A significant difference in these specifications would therefore suggest that the alternative tests are more or less stringent than the traditional MTEL.

Ultimately, the second and third comparisons should be read in conjunction. Comparing those passing the pilot assessments to those passing the traditional MTEL is the best comparison of how candidates accessing the profession through the pilots compare to traditional methods used in Massachusetts. But they describe the participants as an entire group, not the subset of candidates whose employment was changed by participation in the pilot. Adjusting for differences in candidates’ backgrounds provides a test of whether this subset of candidates differs from the larger group of pilot participants and is informative about the relative stringency of the alternatives. As we show below, our findings are generally not sensitive to the choice of comparison.

### Communications and Literacy Skills Alternatives

In Exhibit 30, we show estimates of the relative effectiveness of teachers who passed one of the CLST alternatives. The three specifications correspond to the comparisons we describe above. The first comparison (Experience Only) only adjusts for prior teaching experience and compares teachers passing CLST alternatives to other educators in Massachusetts. The second comparison (MTEL Pilot) compares teachers passing CLST alternatives to teachers passing the traditional CLSTs during the pilot window. The final comparison (Prior CLST Results) adjusts for prior testing history by comparing teachers who passed a CLST alternative to teachers who had similar prior testing histories on the CLSTs.

Exhibit 30. Effectiveness of Candidates Passing CLST Alternatives



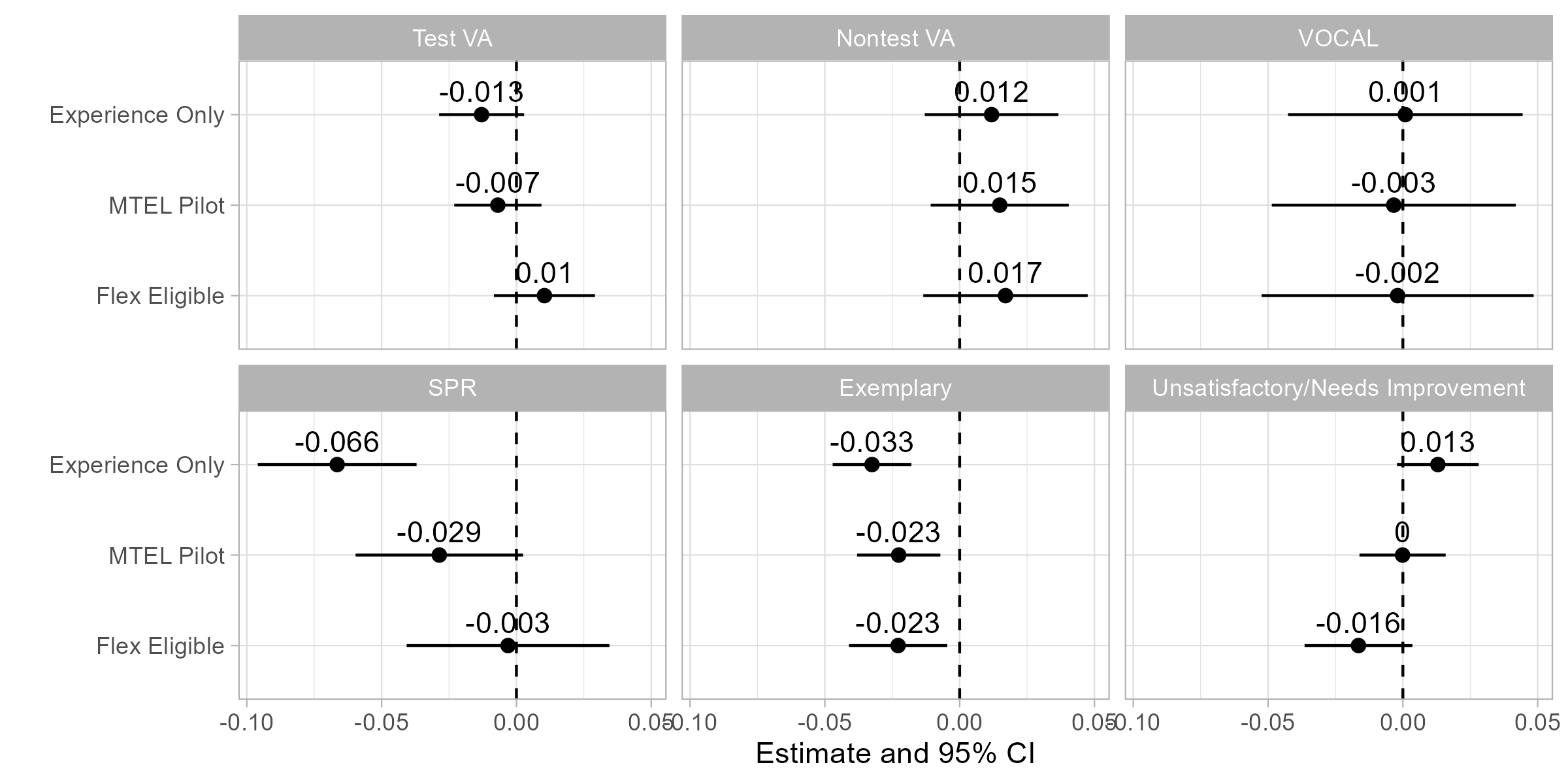
Overall, teachers passing the CLST alternatives performed similarly on a composite test value-added (VA) outcome as other teachers in Massachusetts. The coefficients are not statistically significant and are not sensitive to the choice of comparison group. However, this result does mask some heterogeneity by subject area (see Appendix B). The negative point estimates for the “Test VA” plot are driven by teachers in math. Those passing the CLST alternatives tended to be less effective at raising math achievement than those passing the traditional MTEL CLST, although the difference is only marginally significant. By contrast, there was little difference in teacher value-added on ELA tests. Candidates passing a CLST alternative performed worse on nontest VA (absences, suspensions, grades, and on-time progression), although the differences are not statistically significant. Their students also reported lower scores on the VOCAL survey of school climate. The differences are only significant when compared to all other teachers in the state (Experience Only), although the estimates are similar across the comparison groups.

In the second row of Exhibit 30, we consider performance evaluations. Teachers passing a CLST alternative had summative performance ratings (SPR) that were about 0.07 standard deviations below the average teacher with similar experience. However, this appears to reflect the composition of teachers sitting for licensure exams during the pilot. Compared to other teachers passing the traditional MTEL, the difference is only 0.02 standard deviations and not statistically significant. In the remaining plots, we show that CLST pilot participants generally received similar ratings as other teachers. The significant difference on overall performance ratings when compared to all other teachers appears to be driven by higher rates of unsatisfactory or needs improvement ratings (2 percentage points) and lower rates of exemplary performance (2 percentage points), but neither of these differences is significant.

### MTEL-Flex

In Exhibit 31, we compare teachers passing the MTEL-Flex to other educators. The three specifications correspond to the comparisons we describe above. The first comparison (Experience Only) only adjusts for prior teaching experience and compares teachers passing MTEL-Flex to other educators in Massachusetts. The second comparison (MTEL Pilot) compares teachers passing MTEL-Flex to teachers passing one of the corresponding MTEL subject tests during the pilot window. The final comparison (Flex-Eligible) adjusts for prior testing history by comparing teachers who passed MTEL-Flex to teachers who qualified for MTEL-Flex but passed the traditional MTEL instead.

Exhibit 31. Effectiveness of Candidates Passing MTEL-Flex



Based on the student outcomes, we do not find evidence that teachers who participated in the MTEL-Flex pilot were less effective than other teachers in Massachusetts. Teachers passing MTEL-Flex had similar test and nontest VA and their students reported similar school climate. This is true whether we compare teachers to all other educators, to other teachers passing traditional MTEL subject tests during the pilot, or to other educators who qualified for the MTEL-Flex but passed the traditional MTEL instead.

In Appendix B, we disaggregate results for MTEL-Flex by subject. There is little evidence of any difference in effectiveness between teachers passing MTEL-Flex and those passing the traditional MTEL in either math or ELA. The positive point estimate for the “Flex Eligible” comparison in the “Test VA” plot is driven by math, which suggests that those taking the MTEL-Flex tended to outperform similar teachers in math classrooms. We do, however, find some evidence that those passing the MTEL-Flex produced smaller achievement gains on science tests.

We do find some differences in teachers’ performance evaluations. Teachers who passed the MTEL-Flex had overall performance ratings that were about 0.07 standard deviations below the average teacher with similar levels of experience. Part of this appears to be due to the types of teachers taking subject tests during the pilot. Due to the emergency license provisions that coincide with the pilot, many teachers were taking subject tests later in their careers than is typical. Furthermore, this group of teachers appears to be less effective on average than newly licensed teachers in Massachusetts (Backes et al., 2024b). When we limit the comparison to others passing subject tests in the pilot period, the difference drops to 0.03 standard deviations.

One potential explanation for this finding is that MTEL-Flex participants were required to have attempted and failed the MTEL. The weaker performance evaluations could reflect either differences in who was eligible for the MTEL-Flex or it could indicate that the relatively higher pass rates on the MTEL-Flex qualified less effective teachers for licensure. When we limit the comparison to those whose prior test scores on the traditional MTEL would have qualified them for the MTEL-Flex, we find little evidence of any difference in performance ratings. This finding suggests that the newly-employed candidates performed similarly as those who entered the profession after initially failing the MTEL but subsequently passing a latter attempt.

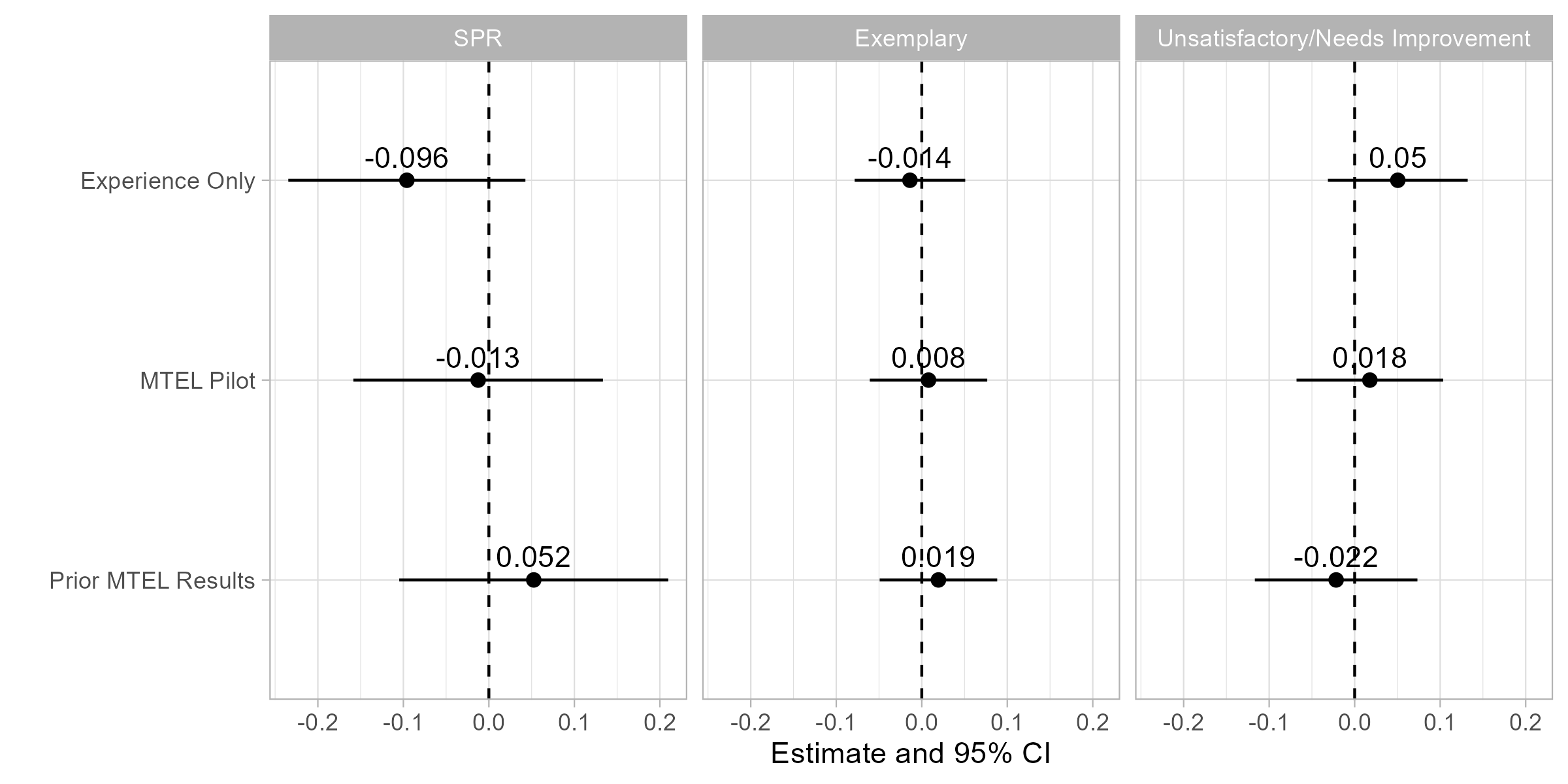
Finally, we see in the remaining two plots that the differences in performance ratings are driven by lower proportions of MTEL-Flex passers earning exemplary ratings (roughly the top 10-15% of educators in Massachusetts). Notably, when we examine teachers who selected into teaching on the basis of their performance on the MTEL-Flex (Flex-Eligible), we find some evidence that this group performed closer to the mean; these teachers were both less likely to earn an exemplary rating and less likely to earn an unsatisfactory or needs improvement rating.

### Program Attestation

In Exhibit 32, we compare performance evaluations for teachers successfully completing the program attestation process to other educators. Due to the limited number of participants in the program attestation option, we are unable to assess their contributions to student outcomes.

The three specifications in Exhibit 32 correspond to the comparisons we describe above. The first comparison (Experience Only) only adjusts for prior teaching experience and compares teachers passing through program attestation to other educators in Massachusetts. The second comparison (MTEL Pilot) compares teachers passing the attestation option to teachers from the same preparation programs who pass one of the corresponding MTEL subject tests during the pilot window. The final comparison (Prior MTEL Results) adjusts for prior testing history by comparing teachers who passed through attestation to teachers from the same programs who had previously failed one of the subject MTEL.

Exhibit 32. Effectiveness of Candidates Passing Program Attestation



Although the estimates are imprecise due to the small number of participants, we do not find significant differences in performance ratings between candidates participating in program attestation and other educators. When compared to all other educators with similar experience levels, the point estimates are negative (but insignificant), with the apparent difference being driven by higher proportions of unsatisfactory or needs improvement ratings. These differences mostly disappear when we limit the comparison to others passing tests from the same programs and are positive (albeit insignificant) when we adjust for prior testing history. We caution, however, that the estimates are quite imprecise. We cannot rule out differences of about 5 percentage points on the likelihood of receiving an exemplary rating or about 10 percentage points on the likelihood of receiving an unsatisfactory or needs improvement rating.

## Summary

The MTEL pilot alternatives were launched between February 2021 through January 2022. In this report, we assess participation through June 2024. As of June 30, 2024, a total of 3,947 teacher candidates had taken at least one of the pilot assessments. The participants were more diverse than other Massachusetts educators in terms of ethnoracial identity and preparation background. Across all the alternative assessments, those who passed alternative assessments were more likely than those passing traditional assessments to be working as Massachusetts teachers on emergency licenses. They were also more ethnoracially diverse than teachers using the traditional MTEL to meet licensure requirements.

Access to the profession, as measured using pass rates, varied across the alternative assessments. The overall pass rate across the CLST alternatives was similar on average to the traditional MTEL, although there were meaningful differences between the different tests being administered. Pass rates on the MTEL-Flex were higher than on the traditional MTEL subject tests, and they have increased over time. MTEL-Flex participants were about 16 percentage points more likely to pass the test than those retaking the traditional MTEL. Program attestation pass rates were also significantly higher than the traditional MTEL. We estimate differences in pass rates of about 29 percentage points after accounting for candidates’ backgrounds and prior test history.

The higher pass rates on the MTEL-Flex and program attestation options appear to have translated into higher rates of licensure and employment. We find that candidates taking the MTEL-Flex were about 2-3 percentage points more likely to have earned an initial or provisional license or to have become employed as teachers. We estimate that program attestation increased employment rates by about 8 percentage points.

We find limited evidence of differences in teacher effectiveness between pilot participants and others passing the traditional MTEL. Candidates passing the CLST alternatives had similar contributions to student achievement and performance evaluations as those passing the traditional MTEL, although their students did report weaker classroom climate and there is suggestive evidence of lower math achievement. Candidates passing the MTEL-Flex performed similarly on value-added metrics as those passing the traditional MTEL, although we do find that they were less likely to earn the top performance evaluation rating in Massachusetts. We do not find any statistically significant differences in performance evaluations between participants in the program attestation option and other teachers, although the sample of teachers is not large enough to rule out meaningful differences.

## Considerations

In this report, we assess the alternative assessments as they have been implemented. As the state considers whether to continue offering these assessments, it is important to note several caveats about generalizing the results beyond the bounds of the pilot.

First, our analyses of teacher effectiveness compare teachers who completed licensure requirements using the piloted assessments to those using the traditional MTEL. We find limited evidence that pilot participants differed from other teachers in their contributions to student outcomes or performance evaluations. Although the fact that these teachers were similarly effective as other educators suggests that the pilot assessments would have limited effects on student learning, these analyses may understate the overall effects of increasing the supply of teachers. The MTEL-Flex and program attestation options appear to have expanded teacher supply by increasing the number of teachers passing licensure tests. Given that some schools are currently struggling to hire qualified educators, improvements in teacher supply might improve hiring outcomes or lead schools to reduce class sizes. These effects would not be captured by the analyses presented in this report.

Second, the difficulty of the assessments might change if they were to become permanent. Pass rates on the MTEL-Flex fluctuated over the course of the pilot. Although we have limited evidence on explanations for shifts in pass rates, the pilot coincided with significant changes in academic performance and study habits, including disruptions to student learning during the pandemic (Goldhaber et al., 2023) and the increasing use of generative AI tools in academic and professional work (Bick et al., 2024; Sidotti & Gottfried, 2023; Singer, 2023). We also might expect pass rates to continue to rise if teacher candidates or preparation programs become more knowledgeable about an assessment and more effective at preparing for it. Similarly, the program attestation option provides preparation providers with considerable flexibility over assessment design. If prospective students base their decisions about where to enroll in part on completion rates, programs may face incentives to increase pass rates on the attestation option. Competition for students might lead programs to compete on leniency (e.g. Mast, 2020). Standardizing assessment materials may help to ensure consistent standards across programs.

Finally, the composition of the pool of test-takers may change as the assessment options become more widely known. Faculty from the preparation providers offering attestation noted that it took some time for candidates to understand their eligibility for this option. This observation is consistent with enrollment patterns for the CLST alternatives and MTEL-Flex. In all cases, early adopters may differ from those who would take the assessments if they were made permanent. If interest in these alternative assessments continues to expand, the characteristics of teachers using them to enter the profession may also change. It is difficult to assess *a priori* how increased participation would influence the supply of teachers. For instance, we find some evidence that the earliest MTEL-Flex participants performed better on the test than later participants. This might suggest that pilot participants would be more effective than those who would take the assessments after the pilot period ends. On the other hand, the pilot assessments might attract candidates with more diverse backgrounds as their availability becomes more widely known. If the alternatives were continued, it may be necessary to continually monitor them to ensure that any alternative maintains adherence to the objective of the assessment program overall.

1. Program Attestation Interviews

We conducted interviews with 11 respondents who are either program coordinators of the program attestation subject matter knowledge (SMK) alternative assessments or content experts responsible for designing and/or administering the alternative assessments. In all interviews, we followed a structured protocol about the process of developing and implementing the SMK attestation. We probed about what is working well with the assessment and where there might be challenges. We also asked about their interest in continuing and potentially expanding their alternative assessments. All interviews were recorded, transcribed, and analyzed for common and divergent themes across different educator preparation programs and subject tests.

Development of the Program Attestation

Program coordinators and content experts describe a few considerations in deciding to develop a program attestation as an SMK alternative assessment. Most people report the process was driven by individual faculty or several faculty after the call for applications was issued. Two programs provided stipends to the faculty who developed the attestations. A few respondents report their decision-making process as considering which MTEL they saw students struggling to pass and deciding to develop an alternative. In some cases, one person created the attestation, while other attestations were created by teams of faculty. To date, only one program has developed an attestation for the General Curriculum Test #2 (Mathematics, Science, and Technology/Engineering) MTEL, which is unique in that it crosses multiple departments (math, science, and engineering/technology) within the educator preparation program, and within science, requires the content expertise of multiple science faculty. Respondents report that this attestation required additional coordination and necessitated multiple people working together across departments, making it more complicated than other attestations to both develop and administer.

In designing the program attestation, faculty typically report starting with the MTEL itself, although one person reported starting with the SMK standards developed by DESE. Some faculty did an analysis of published MTEL questions, and other faculty focused more on the objectives. For one program attestation, faculty developed a diagnostic assessment that students take.

Structure of the SMK Alternative Assessment

The structure of the program attestation varies across programs and subject areas. In three educator preparation programs, students only demonstrate mastery in the subareas in which they did not have an adequate score on the traditional MTEL. In contrast, at two programs, students must demonstrate mastery in all objectives of the MTEL. Students typically have the opportunity to resubmit materials or retake an attestation test or section of an attestation test if they do not pass the alternative assessment the first time.

The program attestations vary across programs in how students demonstrate mastery. The educator preparation provider that offers an attestation for the General Curriculum Test 2 requires that students take a test to demonstrate their mastery in areas where they did not show proficiency on the traditional MTEL. Four programs offer an alternative SMK assessment for the Foundations of Reading MTEL. At one program, students take a test that the faculty describes as very similar to the MTEL, with multiple choice questions and a few open responses. Another program requires students to take a test that was developed by a reading specialist at the school, and faculty believe it is more difficult than the traditional MTEL. At a third program, students watch modules that correspond with the objectives needed based on their diagnostic assessment. They must submit their notes and also score 100% on the quiz for each module they are required to complete. A fourth program created 3 tasks and a final project that encompass the objectives of the Foundations of Reading MTEL. Programs report using rubrics to assess the students’ work.

For the program attestation of the ESL MTEL, at one program, students present a portfolio with assignments that meet each of the objectives in the subareas where they are working toward demonstrating mastery. They may use past assignments if they meet the standards of the rubric, and several subareas require showing new work. For instance, one assignment is a video where they assess a sample of student work. At a second program, students participate in a zero-credit course that includes 5 assignments that students can do on their own and 3 assignments that they do with the faculty member on zoom. These assignments correspond with the 10 objectives of the MTEL. Students can submit previous assignments if it meets the standards of the rubric. In this program, students record themselves teaching. At a third program, students start by doing a self-assessment to identify strengths and areas for growth to improve their practice. Students are required to document their mastery in each area. The program then works with them to create strategies for how they are going to develop the areas of need, which might include observations, reading, preparing specific lesson plans. Then, as part of the assessment, the student is observed. The student writes the rationale and full lesson plan to show what content they have included. After the observation, the faculty provides feedback to the student. Faculty also report using rubrics to assess students’ mastery.

Criteria for Participation

The criteria for participating in the program attestation varies across educator preparation programs. Different criteria across all programs and assessments include the following conditions:

* + Attempted the MTEL;
  + Earning a specific MTEL score;
  + Earning a specific GPA; and
  + Having taken specific courses.

*Attempted the MTEL*: In four of the five programs, students must take and not pass the specific MTEL test. One program has a different set of criteria that does not require taking the traditional MTEL. At that program, students must meet one of the following criteria: 1) Take and not pass the MTEL (no required score); or 2) Have a documented accommodation or IEP; or 3) First language is not English; or 4) Undergraduate GPA is less than 3.0.

*MTEL Scores*: For the four programs that require that students take and not pass the MTEL, three programs require the student earn a specific score on the MTEL. The required score is either 220 or 225, depending on the program. One program does not require students to earn a specific score on the MTEL.

*GPA:* Three programs require a 2.8 or 3.0 GPA to participate.

*Courses Taken:* Two programs require that students have taken certain courses to be eligible. At one program, they require students to have taken the MTEL prep course and earned a B. Another program requires that students have completed courses that cover the objectives of the MTEL.

*Student status*: Four of the five programs require participants to be current students, while one program also allows former students of their program to participate in the alternative assessment.

Several respondents report that it has taken a little bit of time for students to become aware of the alternative assessment opportunity, despite their efforts to get the word out. They have used newsletters, websites, flyers, and the advising process to reach out to students. Multiple people said that interest has been growing as the information about the opportunity has spread and more students are aware of the alternative assessment.

Preparation and Support of Students

Preparation for and support of students seeking to take the program attestation varies across programs and subject tests. Several programs offer courses to help students prepare for the assessment. One program has a specific prep course that is required. One program offers modules for students to watch based on their areas of need. Program administrators report spending a lot of time helping to individualize support for each student. For instance, one program coordinator provides individualized binders for each student and the program also provides coaching. Another program offers 1:1 help in addition to the course. Programs also provide study guides and resources for materials. Respondents reported that they have found the preparation and support sufficient for students to be successful on the alternative assessment. The length of time that students spend preparing for the assessment varies based on how many areas they are working on and how much effort they put into their preparation. Faculty reported students typically spending several months preparing for the attestation.

Diversification of the Workforce

Respondents share mixed responses about whether the program attestation has helped increase diversity of teacher candidates. A few people replied that they do not know, since people’s identities are not always known or visible. One respondent stated that it is “too soon to tell.” Another respondent agreed, explaining that “with more outreach, more exposure, more consistency of the alternative, it has strong potential to help diversify the teacher workforce.” Several respondents, however, reported a belief that the ESL alternative assessment has helped to diversify the workforce. Respondents describe the diversity in terms of increasing the number of teachers of color as well as the number of nontraditional candidates, in terms of age and people whose first language is not English. One respondent also described a candidate who is hearing impaired who was able to gain licensure through the SMK alternative and is now teaching in a middle school.

Ensuring Rigor of the Program Attestation

All respondents reported that they believe their alternative assessment either meets or exceeds the expectations and rigor of the traditional MTEL. Their general reasons include the perceptions that the assessment was developed by content experts at the participating preparation providers; it covers the range of topics on the MTEL; it includes more difficult questions/assignments; and/or it requires candidates to provide substantial evidence. One program coordinator explained that the “MTEL is one measure on one day – whereas this [alternative assessment] can extend over a semester or even of year of continuing to build that knowledge towards competency and meeting those objectives. It is more of building of that knowledge rather than one point in time.”

Beyond these general statements, respondents offer more specific reasons for their views. Several people explained that the students had to apply their knowledge. For some of the alternative assessments, students submit videos of themselves teaching with their own critical assessment of their teaching. Other candidates are observed while teaching and provided with specific feedback. A faculty describes students as applying their knowledge within the context of being a teacher. Another faculty explains that candidates are doing real applications and that their assignments are from their real life. For example, some tasks include: students developing 2 days of instruction that address all 5 literacy domains for 2 different groups of students or selecting assessment data from the school they are working at and submitting a report that describes what the student needs are, strengths, and instructional recommendations. A faculty states they believe that these types of assignments make the assessment more meaningful.

Perception of Success of the Program Attestation

All respondents readily listed different aspects of the SMK alternative assessment that they believe is working well. At the core, multiple respondents report that the SMK alternative assessment has enabled people who they believe are and will be strong teachers to get licensed. A faculty explains that the alternative assessment “has been successful with getting really strong teaching candidates who have a really hard time with taking tests to earn a license.” Similarly, a program coordinator states that they have “wonderful students who have subject matter knowledge who didn’t do well on MTEL.” This belief was expressed by other respondents as well. A few people also shared that they have seen students pass the MTEL who are not necessarily strong teachers. A respondent states that the “MTEL is not the best avenue for demonstrating their capacity as a teacher.” Another person echoes that view, explaining that “MTEL does not tell us about their real-life performance.”[[14]](#footnote-15) These faculty and program coordinators believe the alternative assessment is an important opportunity to enable people who they believe are strong teachers become licensed teachers.

Multiple respondents also report that the fact the alternative assessment is offered free-of-charge for the teacher candidates as an important advantage, making it more equitable and financially accessible to those candidates who are already enrolled in an educator preparation program. Depending on the license the teacher is seeking, candidates are often required to take multiple MTEL exams, which is expensive. Some respondents shared that, depending on where the students live, the availability and location of the traditional MTEL can be a challenge. Students may be required to drive and to travel a fair distance to get to a testing center.

Other respondents describe the alternative assessment as supportive of students. One faculty explains that “Not everybody shows what they know in the same way.” Students are taking the assessment in a familiar environment with faculty who are invested in their growth. It helps lower stress and test anxiety. Depending on the structure of the specific assessment, students can do the alternative on their own timeline. At one program, students can begin their practicum sooner. In this program, students can only start their practicum after they have passed all but one of the required MTELs.[[15]](#footnote-16) Before the SMK alternative assessment, it was taking students longer to participate in the practicum. For a variety of reasons, ranging from financial accessibility to better supporting students to faculty members’ perception that it is enabling people who are good teachers get licensed, all respondents believe the program attestation has been successful.

Challenges with the Program Attestation

The three main challenges that respondents identify are the lack of funding to support this initiative; some of the bureaucratic requirements, either getting past assignments from other professors or in uploading the data to the Department of Elementary and Secondary Education (DESE); and getting the program up and running. At three of the educator preparation providers, the content experts are running the SMK alternative program with no extra pay or release from teaching. They also developed the assessment without any compensation for their time. At the other two programs, the university has provided a stipend and some compensation to the faculty who are running the program. While all respondents expressed satisfaction with the program and the work, they reported the work to develop and run the program as time consuming on top of their teaching requirements and other responsibilities. With regards to the bureaucratic challenges, at one program, a respondent reported difficulties with “trying to get the assignments from past instructors” and also the time-consuming task of explaining the directions to candidates about the requirements for demonstrating mastery in each objective. Another respondent described a lot of difficulties in submitting scores to DESE. Finally, a few respondents explained that it was challenging to get the program up and running and get the word out to eligible students about this opportunity, although they now think students are mostly aware of the opportunity.

Interest in Continuing and/or Expanding Beyond the Pilot

All 11 people interviewed enthusiastically stated interest in continuing with the Program Attestation, using words such as “definitely” and “absolutely.” A few respondents reported that if the program continues, they would like to make some changes to their current attestations. In terms of their interest in developing additional assessments, respondents were less certain. One person stated that it would depend on whether the Department would provide any supports. Another respondent replied that it would depend on if their colleagues in other departments had interest in creating an assessment.

All respondents thought it would be possible and potentially positive to allow teacher candidates outside of their preparation program participate in the alternative assessment program. At the same time, they clearly stated that a number of details would need to be determined to make this idea feasible. Some mentioned the need to figure out some of the logistical pieces, such as if a course is required to participate, how they would accommodate people outside of their programs. A few programs reported already being approached by either individuals or districts interested in participating in the alternative assessment. Currently, all programs are offering the alternative assessment to their students free-of-cost. Given the current time commitment and individualized approach to running the program, multiple people stated they would need more funding and capacity if they were to expand to students beyond their current program.

Changes Recommended by Faculty

Several respondents had some suggested changes for the program, if it continues. One respondent suggested it would be helpful to have an annual window for revisions. A few people reported that they would like to make some adjustments to their alternative assessments. A more general suggestion relates to the structure of the alternative assessment statewide. A few respondents suggested creating a more centralized and consistent approach across educator preparation programs. One respondent stated that the program should not be institution-specific. They suggested the state creating a repository of exams and/or exam questions that programs could tailor to the needs of their students. Similarly, a different respondent stated that the institutions “should be on the same page and agree to the same criteria” and that “a common streamlined process needs to be outlined and be very clear.” Another respondent suggesting having “all the programs collaborate to create the alternative assessments” and that “having consistent attestations and consistent pathways is really important.” In addition, several respondents discussed the importance of having some financial supports for these efforts.

1. Additional Results

Pass Rates by Teacher Candidate Ethnoracial Identity

We summarize differences in pass rates across the alternative assessments in the “Access to the Profession” Section. In this appendix, we list pass rates by candidate ethnoracial identities on the CLST Alternatives and the MTEL-Flex. We suggest some caution interpreting differences across groups of candidates given that the number of participants can become small in some cases. For comparison, we provide the difference in pass rates between each test for each ethnoracial identity group.

* 1. Teacher Candidate Pass Rates on CLST Alternatives by Ethnoracial Identity

|  |  |  |  |
| --- | --- | --- | --- |
| **Candidate Race/Ethnicity** | **Pass Rate on MTEL CLST** | **Pass Rate on CLST Alternatives** | **Difference in Pass Rates** |
| Black | 48.1% | 37.7% | -10.5 |
| Asian or Pacific Islander | 73.0% | 59.4% | -13.6 |
| Hispanic | 55.6% | 45.1% | -10.4 |
| White | 76.0% | 66.2% | -9.8 |
| Other | 65.8% | 56.8% | -9.0 |

Overall, pass rates on the CLST alternatives are about 13 percentage points lower than on the traditional MTEL CLST.[[16]](#footnote-17) The differences in pass rates are relatively constant across candidates’ ethnoracial identities. For the groups in Exhibit B-1, the difference in pass rates is consistently about 10 percentage points. The one outlier is Asian or Pacific Islander candidates, whose pass rates are about 14 percentage points lower on the CLST alternatives.

* 1. Teacher Candidate Pass Rates on MTEL-Flex by Ethnoracial Identity

|  |  |  |  |
| --- | --- | --- | --- |
| **Candidate Race/Ethnicity** | **Pass Rate on Traditional MTEL** | **Pass Rate on MTEL-Flex** | **Difference in Pass Rates** |
| Black | 39.9% | 62.6% | 22.7 |
| Asian or Pacific Islander | 45.1% | 82.5% | 37.5 |
| Hispanic | 51.4% | 74.6% | 23.2 |
| White | 53.6% | 67.6% | 14.0 |
| Other | 50.0% | 82.1% | 32.1 |

We present a similar disaggregation in Exhibit B-2 for the MTEL-Flex. The sample includes those who have taken an MTEL subject test with a score that qualifies them for MTEL-Flex and who elect to retake either the traditional MTEL or the MTEL-Flex. For the MTEL-Flex, there are meaningful differences in pass rates across ethnoracial identify groups. Asian or Pacific Islander candidates, Hispanic candidates, and Black candidates all perform significantly better on the MTEL-Flex.

Finally, in Exhibit B-3, we show reattempt and pass rates for MTEL subject tests by ethnoracial group. We use the sample who scored between 0 and 2 standard errors of measurement below the minimum qualifying score analyses in the MTEL-Flex Qualification and Test Resubmissions section (see Exhibits 24 – 26). Overall, about 77% of candidates in this sample retake either the MTEL or MTEL-Flex. However, the pass rates differ significantly across groups: retesting rates are 78% for white candidates and 79% for Asian American candidates; by contrast, 71% of Hispanic candidates and 66% of Black candidates retake the test.

The differences in retesting rates are mirrored in the proportion of candidates who retake the test and whose next attempt results in passing score. Among white or Asian American candidates with a failing score, 36% and 38%, respectively, retake the test and pass on the next attempt. These rates are 30% and 25%, respectively, for Hispanic and Black candidates.

* 1. Teacher Candidate Pass Rates on MTEL-Flex by Ethnoracial Identity

|  |  |  |
| --- | --- | --- |
| **Candidate Race/Ethnicity** | ***Percentage of MTEL test-takers with below-passing scores who...*** | |
| ***...retake the MTEL or MTEL-Flex*** | ***...retake the MTEL or MTEL-Flex and pass on next attempt*** |
| Black | 66.3% | 25.3% |
| Asian or Pacific Islander | 79.0% | 37.6% |
| Hispanic | 71.2% | 29.6% |
| White | 78.1% | 35.6% |
| Other | 78.8% | 29.7% |

Teacher Effectiveness by Subject

In the Teacher Effectiveness section, we estimate the contributions of teachers participating in the MTEL pilots to student achievement on MCAS. We pool observations over tests and subject areas to increase statistical precision. But these comparisons may mask different effects of teachers across licensure areas. In this Appendix, we separately consider teacher effects on math, English language arts (ELA), as well as on science tests. Math and ELA are tested in grades 3 – 8 and 10 in each year. We use data from students in grades 4 – 8 (we must omit 3rd grade so that we can compare teachers whose students had similar academic performance in the prior year) for these subjects. For science, we use data from students in grades 5, 8, 9, and 10 who took a science test.

We show results for the CLST alternatives in Exhibit B-3. For the CLST alternatives, teacher performance is weaker in math and science than in ELA. The results for math are marginally significant when we compare teachers passing the CLST alternatives to those passing the traditional MTEL CLST during the pilot window. The results for science are imprecisely estimated and not significant under any conventional definition. By contrast, the results for ELA are close to zero across the different comparisons.

* 1. Teacher Contributions to Student Achievement by Subject, CLST Alternatives

|  |  |  |  |
| --- | --- | --- | --- |
| Comparison | Math | ELA | Science |
| Experience Only | -0.041\*\* (0.019) | -0.004 (0.015) | -0.047 (0.031) |
| MTEL Alternatives Pilot | -0.033\* (0.020) | -0.004 (0.015) | -0.050 (0.033) |
| Prior CLST Results | -0.031 (0.020) | -0.001 (0.015) | -0.049 (0.032) |
| Observations | 1,275,796 | 1,249,200 | 428,504 |

|  |
| --- |
| *Notes:* Value-added regressions of student test scores on teacher pilot participation status by subject.  \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 |

In Exhibit B-4, we disaggregate teacher value-added by subject for participants in the MTEL-Flex pilot. The estimated contrasts between those passing the MTEL-Flex and others passing the traditional MTEL during the pilot are close to zero for both math and ELA. We do find some evidence that those passing the MTEL-Flex are more effective at raising student achievement than those who qualified for the MTEL-Flex but instead retook and passed the traditional MTEL, although these results are only significant at the 10% level. As with the CLST alternatives, those passing MTEL-Flex produce lower achievement gains in science, although the results are again imprecise.

* 1. Teacher Contributions to Student Achievement by Subject, MTEL-Flex

|  |  |  |  |
| --- | --- | --- | --- |
| Comparison | Math | ELA | Science |
| Experience Only | -0.010 (0.011) | -0.009  (0.010) | -0.038\* (0.020) |
| MTEL Alternatives Pilot | 0.003 (0.011) | -0.010 (0.011) | -0.035\* (0.020) |
| Flex Eligible | 0.024\* (0.013) | 0.005 (0.012) | -0.025 (0.024) |
| Observations | 1,275,796 | 1,249,200 | 428,504 |

*Notes:* Value-added regressions of student test scores on teacher pilot participation status by subject.   
\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

## References

Angrist, J. D., & Guryan, J. (2004). Teacher testing, teacher education, and teacher characteristics. *American Economic Review*, *94*(2), 241–246. <https://doi.org/10.1257/0002828041302172>

Angrist, J. D., & Guryan, J. (2008). Does teacher testing raise teacher quality? Evidence from state certification requirements. *Economics of Education Review*, *27*(5), 483–503. <https://doi.org/10.1016/j.econedurev.2007.03.002>

Bacher-Hicks, A., & Koedel, C. (2023). Estimation and interpretation of teacher value added in research applications. In *Handbook of the Economics of Education* (Vol. 6, pp. 93–134). Elsevier. <https://doi.org/10.1016/bs.hesedu.2022.11.002>

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2022). *Teachers and school climate: Effects on student outcomes and academic disparities* (CALDER Working Paper 274–1022). Center for the Analysis of Longitudinal Data in Education Research.

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2024a). How to measure a teacher: The influence of test and nontest value-added on long-run student outcomes. *Journal of Human Resources*. Advance online publication. <https://doi.org/10.3368/jhr.1023-13180R2>

Backes, B., Cowan, J., Goldhaber, D., & Theobald, R. (2024b). Four years of pandemic-era emergency licenses: retention and effectiveness of emergency-licensed Massachusetts teachers over time. *Economics of Education Review, 101*, 102562. <https://doi.org/10.1016/j.econedurev.2024.102562>

Bennett, C. I., McWhorter, L. M., & Kuykendall, J. A. (2006). Will I ever teach? Latino and African American students’ perspectives on PRAXIS I. *American Educational Research Journal, 43*(3), 531–575. <https://doi.org/10.3102/00028312043003531>

Bick, A., Blandin, A., & Deming, D. (2024). *The Rapid Adoption of Generative AI* (NBER Working Paper No. 32966). National Bureau of Economic Research. <https://doi.org/10.3386/w32966>

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. <https://doi.org/10.1016/j.econedurev.2007.10.002>

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. <https://doi.org/10.3368/jhr.45.3.655>

Cowan, J., Goldhaber, D., Jin, Z., & Theobald, R. (2023). Assessing Licensure Test Performance and Predictive Validity for Different Teacher Subgroups. *American Educational Research Journal*. Advance online publication. <https://doi.org/10.3102/00028312231192365>

Cowan, J., Goldhaber, D., & Theobald, R. (2022). Performance evaluations as a measure of teacher effectiveness when implementation differs: Accounting for variation across classrooms, schools, and districts. *Journal of Research on Educational Effectiveness, 15*(3), 510–531. <https://doi.org/10.1080/19345747.2021.2018747>

Gershenson, S. (2021). *Identifying and producing effective teachers* (EdWorkingPaper No. 21–351). Annenberg Institute at Brown University. <https://www.edworkingpapers.com/ai21-351>

Gitomer, D. H., Brown, T. L., & Bonett, J. (2011). Useful signal or unnecessary obstacle? The role of basic skills tests in teacher preparation. *Journal of Teacher Education*, *62*(5), 431–445. <https://doi.org/10.1177/0022487111412785>

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., Cowan, J., & Theobald, R. (2017). Evaluating Prospective Teachers: Testing the Predictive Validity of the edTPA. *Journal of Teacher Education, 68*(4), 377–393. <https://doi.org/10.1177/0022487117702582>

Goldhaber, D., Gratz, T., & 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. <https://doi.org/10.1016/j.econedurev.2017.09.002>

Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2023). The educational consequences of remote and hybrid instruction during the pandemic. *American Economic Review: Insights*. <https://doi.org/10.1257/aeri.20220180>

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy, 126*(5), 2072–2107.

Mansfield, R. K. (2015). Teacher quality and student inequality. *Journal of Labor Economics, 33*(3), 751–788. <https://doi.org/10.1086/679683>

Massachusetts Department of Elementary and Secondary Education. (2021, October 7). *Pilot of alternative assessments to the MTEL*. <https://www.doe.mass.edu/mtel/alt-assess/default.html>

Massachusetts Department of Elementary and Secondary Education. (2022). *MTEL-Flex assessment handbook* [Version 1.0]. Massachusetts Department of Elementary and Secondary Education.

Mast, E. (2020). Race to the bottom? Local tax break competition and business location. *American Economic Journal: Applied Economics, 12*(1), 288–317. https://doi.org/10.1257/app.20170511

National Council on Teacher Quality. (2020). *Licensure requirements* [national results]. State Teacher Policy Database. [Data set]. <https://www.nctq.org/yearbook/national/Licensure-Requirements-90>

Nettles, M. T., Scatton, L. H., Steinberg, J. H., & Tyler, L. L. (2011). Performance and passing rate differences of African American and White prospective teachers on Praxis examinations: A joint project of the National Education Association (NEA) and Educational Testing Service (ETS). *ETS Research Report Series, 2011*(1), 1-82. <https://doi.org/10.1002/j.2333-8504.2011.tb02244.x>

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics, 130*, 105–119. <https://doi.org/10.1016/j.jpubeco.2015.02.008>

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review Papers & Proceedings, 94*(2), 247–252.

Rucinski, M., & Goodman, J. (2019). *Racial diversity in the teacher pipeline: Evidence from Massachusetts* [Policy brief]*.* Harvard Kennedy School.

Sass, T. R. (2015). Licensure and worker quality: A comparison of alternative routes to teaching. *The Journal of Law and Economics*, *58*(1), 1–35. <https://doi.org/10.1086/682904>

Sidotti, O. & Gottfried, J. (2023, November 16). “About 1 in 5 U.S. teens who’ve heard of ChatGPT have used it for schoolwork.” *Pew Research Center Short Reads*. <https://www.pewresearch.org/short-reads/2023/11/16/about-1-in-5-us-teens-whove-heard-of-chatgpt-have-used-it-for-schoolwork/>

Singer, N. (2023, December 13). “Cheating Fears Over Chatbots Were Overblown, New Research Suggests.” *New York Times*. <https://www.nytimes.com/2023/12/13/technology/chatbot-cheating-schools-students.html>



|  |  |
| --- | --- |
|  |  |
|  | About the American Institutes for Research®  Established in 1946, the American Institutes for Research (AIR) is a nonpartisan, not-for-profit institution that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit [AIR.ORG](https://www.air.org). |
|  | AIR Headquarters  1400 Crystal Drive, 10th Floor  Arlington, VA 22202-3289 +1.202.403.5000 | [AIR.ORG](https://www.air.org/) |
| Notice of Trademark: “American Institutes for Research” and “AIR” are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.  Copyright © 2023 American Institutes for Research. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on [AIR.ORG](http://www.air.org/). | |
| XXXXX\_MO/23 | |

1. Employment data for the start of the 2024-25 school year were not available as of the time of this report and were not included. Similarly, teacher performance evaluation data for the 2023-24 school year were not yet available. Student achievement data for the MCAS is typically available earlier than teacher performance evaluations; hence, teacher value added to student test scores and school climate are the only available performance measure for the 2023-24 school year in this report. [↑](#footnote-ref-2)
2. These licensure requirements were temporarily relaxed during the COVID-19 pandemic. Through November 2023, DESE offered an emergency license option that temporarily exempted teachers from completing certain licensure requirements. To continue teaching on a license after their emergency license expires, teachers must advance to an initial or provisional license by completing the relevant licensure tests. [↑](#footnote-ref-3)
3. The standard error of measurement is a measure of how much a candidate’s score would be expected to vary over subsequent administrations of the test due to random chance. [↑](#footnote-ref-4)
4. The state only collects data on completion of an out-of-state preparation program when candidates submit an application for licensure. [↑](#footnote-ref-5)
5. The controls include indicators for whether the teacher has an emergency or initial/provisional license, employment as a teacher, employment in a non-teaching role, prior completion of a preparation program, the number of prior attempts on the CLST, race/ethnicity, gender, and the subtest taken. We also include a quadratic polynomial in the number of days since the beginning of the pilot to account for the time available to reach key milestones. [↑](#footnote-ref-6)
6. In Exhibit B-1, we disaggregate pass rate data by teacher candidate race/ethnicity. The difference in pass rates between the traditional MTEL CLST and the CLST alternatives is relatively constant across ethnoracial identify groups. The results in these analyses may differ somewhat from pass rates published by DESE because we conduct additional identity matching using other administrative records to capture demographic information for as many candidates as possible. [↑](#footnote-ref-7)
7. Ideally, we could test whether effects on retesting rates differ by candidates’ ethnoracial identification. Unfortunately, the limited samples for non-white teacher candidates preclude such analyses. [↑](#footnote-ref-8)
8. For each candidate participating in the attestation pilot, we identify the closest match in terms of prior testing experiences, baseline employment and licensure, testing date, and preparation program using a Mahalanobis distance metric. We use matching rather than covariate adjustment given the small number of participants in the pilot, but the results are similar using ordinary least squares (OLS) regression instead. [↑](#footnote-ref-9)
9. Course grades are only available for students in grades 7 – 8 and 10. We omit course grades for students in grades 4 – 6. [↑](#footnote-ref-10)
10. We estimate a graded response model using the annual performance evaluation data (Cowan et al., 2022). The graded response model is akin to standardizing teacher performance ratings, but more flexibly models teachers with high and low evaluations. [↑](#footnote-ref-11)
11. Teachers are significantly more likely to earn a low rating (unsatisfactory or needs improvement) and significantly less likely to earn an exemplary rating in their first year in a new school (Cowan et al., 2022). [↑](#footnote-ref-12)
12. We also can only compare teachers with comparable outcomes, so, for instance, are limited in the case of value-added effectiveness to comparisons between teachers whose students are in tested grades and subjects. [↑](#footnote-ref-13)
13. Consider a hypothetical experimental evaluation of the pilot assessments where candidates are randomly assigned to take either the traditional MTEL or an alternative assessment and candidates passing either assessment are eligible for employment. If we assume that the assessments themselves do not improve the performance of individual teachers, then any differences in average effectiveness among employed teachers by randomization status would necessarily be due to the movement of the newly-employed group into the workforce. Because we cannot randomize teachers to test type, we instead use prior test performance to adjust for differences in the composition of the two samples of test-takers. [↑](#footnote-ref-14)
14. The MTEL is designed to measure content knowledge and/or communications and literacy skills and does not necessarily measure pedagogical skills or pedagogical content knowledge. The Massachusetts educator licensure system includes other assessment tools, such as the Candidate Assessment of Performance (CAP) and Performance Review Program for Initial Licensure (PRPIL), that are intended to measure pedagogical skill. [↑](#footnote-ref-15)
15. DESE recommends against conditional program enrollment or practicum completion on passing required MTEL, but some programs may implement these requirements. [↑](#footnote-ref-16)
16. The omitted groups from Exhibit B-1, those without any reported race or ethnicity, tend to have lower pass rates. [↑](#footnote-ref-17)