

The Impact of Removing MCAS Graduation Requirement on Student Outcomes

Matthew Krasnow

`mkrasnow@college.harvard.edu`

Russell O'Donnell

`russellodonnell@college.harvard.edu`

May 2, 2025

1 Introduction and Motivation

The Massachusetts Comprehensive Assessment System (MCAS) has been a cornerstone of educational accountability in the state since its implementation as part of the 1993 Education Reform Act. Until recently, high school students were required to pass the MCAS exams in English Language Arts, Mathematics, and Science to earn a diploma. This requirement stood firm regardless of their other academic achievements. In November 2024, Massachusetts voters approved Question 2, eliminating the MCAS as a graduation requirement and shifting toward coursework-based graduation criteria instead (Boston.com, 2024). This significant policy change provides a unique opportunity. We can now study the effects of high-stakes exit exams on student outcomes.

This research proposal addresses a critical question at the intersection of educational assessment and policy. What is the causal effect of removing the Massachusetts Comprehensive Assessment System (MCAS) high school graduation requirement on student outcomes?

We focus specifically on graduation rates, college enrollment, college retention rates, and workforce participation. The last metric includes both employment status and earnings.

Understanding the impacts of this policy change matters. The national context makes it particularly relevant. Since the early 2000s, states have been moving away from high school exit exams. More than half of states required students to pass exit exams in 2002. That number has steadily declined (Boston Globe, 2024). Despite this trend, rigorous research on the consequences of removing these requirements remains limited. Prior research on the introduction of exit exams found mixed or somewhat negative effects on educational outcomes (Caves and Balestra, 2018). The impacts of elimination, however, remain understudied.

The debate around the MCAS graduation requirement highlights competing perspectives on educational assessment. Advocates for the requirement argued that standardized testing ensures consistent academic standards across districts. They believed it motivates schools to focus on foundational skills. As one proponent noted, the MCAS requirement served as "a lifeline for highest-need students" by identifying struggling students for intervention (Rhode Island Current, 2023). Opponents saw it differently. They contended that the requirement disproportionately affected disadvantaged students. The Massachusetts Teachers Association argued that standardized tests cannot "measure the full scope of skills, knowledge, and competencies" needed for success (Boston.com, 2024).

Research tells a complex story. Brown University scholars found that approximately 1.8% of Massachusetts students who completed all other local graduation requirements failed to pass the MCAS exams. These students received certificates of attainment rather than diplomas (Harvard Graduate School of Education, 2024). This affects both their immediate educational attainment and potentially their long-term economic prospects. The research also found that students with higher MCAS scores tend to have better long-term outcomes. These include higher college graduation rates and earnings by age 30. The causal relationship remains unclear, though, due to the implementation of MCAS alongside other educational reforms (Harvard Graduate School of Education, 2024).

National trends in college enrollment add context to our study. According to recent data, 62.2% of recent white high school graduates enrolled in college in 2024. The figures stand at 59.2% for Black graduates and 55.4% for Hispanic graduates. These statistics highlight persistent disparities (Bureau of Labor Statistics, 2024). Understanding how high-stakes testing requirements affect these patterns is crucial for developing equitable education policies.

This proposal outlines two methodological approaches to address our research question: a randomized controlled trial (RCT) and a quasi-experimental difference-in-differences (DiD) design. By comparing these approaches, we aim to determine the most rigorous and feasible method. Our goal: evaluating the causal impact of removing the MCAS graduation requirement on student outcomes in Massachusetts.

2 Method 1: Randomized Controlled Trial

Our primary research question examines the causal effect of removing the MCAS high school graduation requirement on student outcomes. To address this question with a high degree of internal validity, we propose a cluster-randomized controlled trial at the district level.

2.1 Variables and Data Collection

The treatment variable in our study is binary: whether a student must pass the MCAS to graduate. In the experimental context, this depends on the district’s random assignment to either require MCAS (control) or not (treatment). We will measure several key outcome variables:

1. **High school graduation rates:** The percentage of students who earn a standard diploma within four years of entering high school.
2. **College enrollment rates:** The percentage of high school graduates who enroll in any postsecondary institution within six months of graduation.

3. **Academic achievement:** Performance on other standardized assessments that are not tied to graduation requirements.
4. **Workforce participation:** Employment status and earnings after high school for students who do not pursue higher education.

We'll collect these data from multiple sources. Administrative data from the Massachusetts Department of Elementary and Secondary Education (DESE) will provide information on in-school outcomes such as graduation status and test scores. The National Student Clearinghouse database will provide data on college enrollment patterns. For workforce outcomes, we will use state employment records when available. We'll supplement with targeted surveys to track outcomes not captured in administrative data.

2.2 Randomization and Treatment Assignment

The treatment will be randomized at the school district level rather than at the individual student level. This cluster-level random assignment serves several purposes. It avoids creating inequities within schools. It prevents contamination through peer effects. And it reflects how such a policy would realistically be implemented. We will employ stratified randomization based on factors such as district size, prior graduation rates, and demographics. This ensures balance between treatment and control groups on key characteristics.

To achieve sufficient statistical power, we need to include a substantial number of districts. Massachusetts has approximately 400 school districts. We aim to recruit at least 100 districts for participation, with 50 randomly assigned to each condition.

2.3 Intervention Implementation

Districts assigned to the treatment condition will implement a coursework-based graduation criterion. Students can earn a diploma by successfully completing a set of locally certified courses that meet state standards. Students in these districts will still take the MCAS

as required by state law for accountability purposes. Failing the exam will not prevent graduation if all other requirements are met.

Control districts will maintain the status quo. They will require students to pass the MCAS in English, Math, and Science to receive a diploma, in addition to completing required coursework credits.

2.4 Analytical Approach

Our primary analytical model will be a regression of each outcome on a treatment indicator. We will control for district-level covariates and use robust standard errors clustered at the district level (the unit of randomization). The model will take the form:

$$Y_{i,d} = \beta_0 + \beta_1 \text{Treatment}_d + \gamma X_d + \varepsilon_{i,d} \quad (1)$$

Where $Y_{i,d}$ is the outcome for student i in district d , Treatment_d is a binary indicator for whether district d was assigned to the treatment group, X_d is a vector of district-level covariates measured at baseline, and $\varepsilon_{i,d}$ is the error term.

We will conduct intent-to-treat (ITT) analyses as our primary approach. This means comparing all students in treatment versus control districts regardless of individual compliance. Why? It provides an unbiased estimate of the policy's impact as it would be implemented in practice. We'll also examine heterogeneous effects across student subgroups. We'll focus particularly on students who would be on the margin of passing the MCAS, English language learners, students with disabilities, and economically disadvantaged students.

2.5 Internal Validity Considerations

Randomization ensures high internal validity by eliminating systematic baseline differences between treatment and control districts. Several threats remain, however, that we must address:

1. **Non-compliance:** Students may transfer between districts. This could potentially dilute the treatment effect or bias estimates. We'll mitigate this by performing analyses based on original district assignment (intention-to-treat) and tracking student mobility patterns. If transfer rates are significant, we may use an instrumental variable approach. We would use the random assignment of the student's original district as an instrument for whether they actually faced the MCAS requirement.
2. **Spillovers:** Neighboring control districts might alter their instruction if they know nearby districts removed the exam requirement. Students in mixed environments could be indirectly affected. To minimize spillovers, we will clearly communicate the policy at the district level. We'll monitor for evidence of contamination, such as changes in teaching practices in control schools.

2.6 External Validity Considerations

How generalizable will our findings be? It depends on how similar other contexts are to Massachusetts. Our results will be most applicable to states that currently use or are considering phasing out high school exit exams. Massachusetts is a high-performing state educationally. The effects might differ in states with different baseline achievement levels or accountability structures.

To enhance external validity, we will carefully document the implementation process and context. This allows others to judge how the conditions match their local environment. We'll also collect qualitative data on how districts implement the policy change. This may help explain variation in effects and inform adaptation to other contexts.

3 Method 2: Difference-in-Differences Design

As an alternative to an RCT, we propose a quasi-experimental difference-in-differences (DiD) approach. This method exploits the natural variation in implementation timing across dis-

tricts following the policy change. All districts will eventually transition away from the MCAS requirement. Some may do so earlier than others. This could be due to administrative timelines, local school board decisions, or phased implementation plans. Our approach takes advantage of this variation.

3.1 Key Variables in the DiD Method

The key variables in our DiD approach include:

1. **Treatment variable (Treat):** A binary indicator for whether a district has implemented the removal of the MCAS graduation requirement (1 = requirement removed, 0 = requirement still in place).
2. **Time variable (Post):** A binary indicator for whether the time period is after the policy change was announced/began implementation (1 = post-policy, 0 = pre-policy).
3. **Interaction term ($\text{Treat} \times \text{Post}$):** The DiD estimator capturing the causal effect of removing the MCAS requirement.
4. **Outcome variables:** The same outcome measures used in the RCT design—high school graduation rates, college enrollment rates, college retention rates, and workforce outcomes.

For data sources, we will use the same administrative and survey data described in the RCT proposal: DESE data for in-school outcomes, National Student Clearinghouse for college outcomes, and state employment records or surveys for workforce outcomes.

3.2 Regression Specification

Our primary regression model will be:

$$Y_{ist} = \beta_0 + \beta_1(\text{Treat}_s) + \beta_2(\text{Post}_t) + \beta_3(\text{Treat}_s \times \text{Post}_t) + \gamma X_{ist} + \delta_s + \lambda_t + \varepsilon_{ist} \quad (2)$$

Where:

- Y_{ist} is the outcome for student i in district s at time t
- Treat_s is a binary indicator for whether district s is an early adopter of the policy change
- Post_t is a binary indicator for whether time t is after the policy announcement
- $\text{Treat}_s \times \text{Post}_t$ is the interaction term, where β_3 is our causal estimate of interest
- X_{ist} is a vector of student-level controls (demographics, prior academic performance)
- δ_s represents district fixed effects to control for time-invariant district characteristics
- λ_t represents year fixed effects to control for statewide trends affecting all districts
- ε_{ist} is the error term

We'll cluster standard errors at the district level. This accounts for within-district correlation in outcomes. We'll run the regression separately for each outcome variable.

3.3 Identifying Assumptions

The key identifying assumption for our DiD approach is the parallel trends assumption. In the absence of treatment, the average change in outcomes would have been the same for both early-adopting (treatment) and late-adopting (control) districts. For our context, this means that without the policy change removing the MCAS requirement, trends in graduation rates, college enrollment/retention, and workforce outcomes would have evolved similarly in both groups of districts over time.

Additional assumptions include:

1. **No anticipation effects:** Districts and students do not change their behavior in anticipation of the policy change.
2. **No compositional changes:** The composition of students in treatment and control districts does not change differentially over time.
3. **No concurrent policy changes:** No other policies were implemented that differentially affected treatment and control districts during the study period.

3.4 Evidence for Internal Validity

To provide evidence that our assumptions are satisfied, we will:

1. **Test for parallel pre-trends:** Extend our model to include leads and lags of the treatment. This allows us to visually and statistically examine if trends were parallel before the policy change. We would estimate:

$$Y_{ist} = \beta_0 + \sum_k \beta_k (\text{Treat}_s \times \text{Period}_k) + \delta_s + \lambda_t + \gamma X_{ist} + \varepsilon_{ist} \quad (3)$$

Where Period_k represents time periods before and after the policy change. The coefficients on pre-policy interaction terms should not be statistically significant if parallel trends hold.

2. **Conduct placebo tests:** Run our main specification on outcomes that should not be affected by the MCAS requirement. Attendance rates for younger grades not subject to the requirement provide one example.
3. **Perform event study analysis:** Plot the estimated treatment effects over time. This helps visualize any pre-existing trends or delayed effects.

4. **Implement matching analysis:** Combine DiD with propensity score matching. This ensures comparison between districts with similar observable characteristics.
5. **Test for compositional changes:** Compare student demographics and prior achievement across treatment and control districts before and after the policy change.

3.5 Internal Validity Threats

Despite our methodological approach, several threats to internal validity remain:

1. **Selection bias:** Early-adopting districts may be systematically different from late adopters. They might be more progressive, better resourced, or have different student populations. District fixed effects control for time-invariant characteristics. Dynamic differences could still bias our results. If early adopters were already improving faster than others, our estimate would overstate the policy's effect.
2. **Policy endogeneity:** Districts may choose to adopt the policy change based on expectations about how it will affect their students. Districts with more students struggling with MCAS might adopt earlier. This could lead to an overestimation of graduation rate effects.
3. **Student mobility:** Students might transfer between early-adopting and late-adopting districts based on preferences regarding the MCAS requirement. This would contaminate the treatment and control groups. The direction of bias would depend on which students move.
4. **Implementation heterogeneity:** Districts may implement the policy change differently. This creates noise in our treatment variable and potentially attenuates our estimates.

To address these concerns, we'll conduct robustness checks. These include alternative con-

trol groups (e.g., neighboring states with similar education systems), controlling for district-specific time trends, and analyses that account for student mobility.

3.6 External Validity Considerations

Our research would generalize most directly to states with similar educational contexts and achievement levels as Massachusetts. It would be especially relevant to high-performing states with achievement gaps similar to those in Massachusetts. The findings would particularly apply to states using or considering high-stakes exit exams.

The study speaks to broader policy questions about standardized testing in K-12 education. Does high-stakes testing create barriers to graduation that disproportionately affect certain student populations? Can alternative assessment methods maintain educational quality while improving access? Our research addresses these critical questions.

The results may not generalize well to all contexts. States with significantly different educational systems, demographic compositions, or economic conditions might see different outcomes. Additionally, this research specifically addresses the short-term impacts of removing testing requirements. Longer-term effects on college completion, career trajectories, and skill development would require extended longitudinal studies beyond our current scope.

4 Discussion

Having outlined two methodological approaches to address our research question, we now compare them. Which approach would be preferable to implement?

4.1 Tradeoffs in Internal Validity

The RCT design offers superior internal validity. Random assignment of districts to treatment and control conditions ensures this advantage. Districts in both groups should be similar on both observable and unobservable characteristics. This allows for a clean estima-

tion of the causal effect of removing the MCAS requirement. The main threats to internal validity in an RCT—non-compliance and spillovers—can be addressed. Intent-to-treat analysis and careful monitoring of implementation help mitigate these concerns.

The DiD design relies on the parallel trends assumption. This cannot be directly tested for future outcomes. We can examine pre-treatment trends to provide supporting evidence. Yet a possibility remains: treatment and control districts might have diverged even without the policy change. The DiD approach is more vulnerable to selection bias. Early-adopting districts might systematically differ from late adopters in ways that affect outcome trajectories.

4.2 Tradeoffs in External Validity

For external validity, the DiD approach may have an advantage. It studies the natural rollout of the policy change across districts. We observe how the removal of the MCAS requirement plays out in a real-world context. This provides insights into how the policy works under normal conditions, not the carefully controlled conditions of an RCT.

The RCT might suffer from a Hawthorne effect. Districts might implement the policy differently knowing they are part of a study. Additionally, districts that volunteer to participate in an RCT might not represent all districts in Massachusetts. This limits generalizability.

4.3 Feasibility Considerations

The designs differ substantially in feasibility. The DiD approach has a clear advantage. It leverages existing policy variation without requiring researcher control over implementation. Massachusetts has already voted to remove the MCAS requirement. A DiD approach can adapt to the actual implementation timeline across districts.

The RCT faces significant practical challenges. It would require obtaining buy-in from numerous districts willing to be randomly assigned. Some would maintain the graduation requirement. Others would remove it. This might face resistance from district leaders,

parents, and other stakeholders. The contentious nature of the MCAS requirement amplifies this challenge. Furthermore, now that the state has voted to remove the requirement, it might be legally or politically impossible to randomly assign some districts to maintain it.

4.4 Ethical Considerations

Ethical considerations also favor the DiD approach. The RCT would create a situation where otherwise similar students face different graduation requirements based on random assignment. Some might view this as inequitable. If there is reason to believe the MCAS requirement might harm students' educational prospects, some would consider it unethical to maintain this barrier for control district students.

The DiD approach avoids these ethical concerns. It simply studies natural variation in policy implementation timing. Researcher intervention does not determine which students face which requirements.

4.5 Cost Considerations

From a cost perspective, the DiD approach is likely more efficient. It primarily relies on existing administrative data. It does not require the extensive coordination and monitoring needed for an RCT. The RCT would involve costs related to recruitment, coordination with districts, monitoring compliance, and potentially incentivizing participation. A DiD approach reduces or eliminates these costs.

4.6 Decision on Preferred Method

Based on these considerations, we conclude that the difference-in-differences approach is preferable. The RCT would offer stronger internal validity. Yet the practical, ethical, and political challenges make it infeasible in the current policy context. The DiD approach offers a robust alternative. It can leverage natural variation in implementation timing to estimate

causal effects while remaining practical and ethically sound.

The key challenge with the DiD approach will be addressing potential selection bias. Which districts implement the policy change earlier versus later? To mitigate this, we will carefully document district characteristics and implementation processes. We'll use district fixed effects to control for time-invariant differences. We'll conduct extensive robustness checks with alternative specifications and control groups.

Additionally, we propose supplementing the quantitative DiD analysis with qualitative research. We'll conduct interviews with district leaders, teachers, and students. This helps us better understand the mechanisms through which removing the MCAS requirement affects educational outcomes. This mixed-methods approach provides both causal estimates of the policy's impact and insights into the processes that drive these impacts.

4.7 Implications and Limitations

Our study will provide valuable evidence on the effects of high-stakes exit exams on student outcomes. It contributes to an important policy debate in education. Several limitations should be acknowledged, however. First, even with the DiD approach, causal inference relies on assumptions that cannot be directly tested. Second, the focus on short-term outcomes like graduation rates and immediate college enrollment may miss longer-term effects. College completion or career trajectories might show different patterns. Third, the generalizability of our findings to other states with different educational contexts may be limited.

Despite these limitations, this research will make a significant contribution. We will enhance understanding of how graduation requirements affect student outcomes, particularly for vulnerable student populations. The findings will inform not only Massachusetts' implementation of its new graduation policy. They will also guide other states' decisions about high-stakes testing requirements.

References

- Boston Globe (2024). Most states have extensive graduation requirements. In Massachusetts, it's just the MCAS. Accessed April 26, 2025.
- Boston.com (2024). Yes on 2: MCAS will no longer be high school graduation requirement. Accessed April 26, 2025.
- Bureau of Labor Statistics (2024). College enrollment and work activity of recent high school and college graduates summary - 2024 A01 results. Accessed April 26, 2025.
- Caves, K. and Balestra, S. (2018). The impact of high school exit exams on graduation rates and achievement. *The Journal of Educational Research*, 111(2):186–200.
- Harvard Graduate School of Education (2024). What impact have high school exit exams had in Massachusetts? Accessed April 26, 2025.
- Rhode Island Current (2023). MCAS graduation requirement is not punitive, it's a lifeline for highest-need students in Mass. Accessed April 26, 2025.