

# Can Recognition Improve Results? A First Look at the Massachusetts Student Achievement Award Program

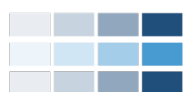
Dan Goldhaber

Zeyu Xu

Tim Daly

October 2026

WORKING PAPER No. 347-1026



**CALDER**

National Center for Analysis of  
Longitudinal Data in Education Research



---

# Can Recognition Improve Results? A First Look at the Massachusetts Student Achievement Award Program

**Dan Goldhaber**

*American Institutes for Research / CALDER*

*University of Washington / CEDR*

**Zeyu Xu**

*American Institutes for Research / CALDER*

**Tim Daly**

*EdNavigator*

---

## Contents

---

|                                                           |     |
|-----------------------------------------------------------|-----|
| Contents.....                                             | i   |
| Acknowledgments .....                                     | ii  |
| Abstract .....                                            | iii |
| 1. Introduction .....                                     | 1   |
| 2. The Massachusetts Student Achievement Awards .....     | 2   |
| 3. Data, Sample, and Empirical Strategy .....             | 5   |
| 4. Findings .....                                         | 7   |
| 5. Discussion and Conclusion.....                         | 11  |
| References .....                                          | 15  |
| Tables and Figures.....                                   | 18  |
| Appendix A. ELA Results .....                             | 27  |
| Appendix B. Results for Top 2% Mathematics Achievers..... | 32  |

## Acknowledgments

---

This work is made possible through the use of confidential data from the Massachusetts Department of Elementary and Secondary Education (DESE). This research was supported by grants from the Carina Fund and the Walton Family Foundation. Funding from Carina launched the project with a focus on the top 2% of students in math, and funding from the Walton Family Foundation enabled expansion of the project to a broader set of outcomes. Together, these resources made the present study possible. The findings and views expressed in this paper are those of the authors and do not necessarily represent DESE, the Carina Fund, the Walton Family Foundation, or the institutions with which the authors are affiliated.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any errors are attributable to the authors.

CALDER • American Institutes for Research  
1400 Crystal Drive 10<sup>th</sup> Floor, Arlington, VA 22202  
202-403-5796 • [www.caldercenter.org](http://www.caldercenter.org)

***Can Recognition Improve Results? A First Look at the Massachusetts Student Achievement Award Program***

Dan Goldhaber, Zeyu Xu, & Tim Daly  
CALDER Working Paper No. 347-1026  
October 2026

**Abstract**

Students from disadvantaged backgrounds are underrepresented among top performers in mathematics and English language arts (ELA) tests, and those who do become top performers are more likely than similarly high-achieving peers to lose ground over time. This study examines the Massachusetts Student Achievement Awards (MSAA) program, which recognized disadvantaged students who had scores judged to be “exceeding expectations” on the Massachusetts state assessment—roughly the top 7% on state standardized tests in mathematics or ELA—with certificates, letters, informational resources, and a modest gift card. Using a difference-in-differences design, we estimate the program's effects on high-achieving, disadvantaged students in Grades 3–7. Recognition increased next-year mathematics test achievement by 0.05–0.07 standard deviations, a result that holds consistently across specifications; effects on ELA achievement are smaller and not consistently significant across model specifications. Within mathematics, effects are larger for girls and for students who were not consistently among the state's top performers. These findings suggest that even among a small population of exceptionally high-performing but under-recognized disadvantaged students, a simple, low-cost, and readily scalable recognition intervention can meaningfully improve short-term academic outcomes—particularly in mathematics.

## 1. Introduction

Disadvantaged high-achieving students are significantly less likely than similar peers to sustain that performance, enroll in advanced coursework, or reach selective colleges (Goldhaber et al., 2021; Hoxby & Avery, 2013). These trends reflect the interaction of many factors. In this study, we examine an intervention that targets one of those factors: information barriers (Page & Scott-Clayton, 2016). Disadvantaged students often underestimate their academic abilities (e.g., Cunha, 2009; Hoxby & Turner, 2015), and their teachers often underestimate them as well (Anderson, 2024; Batruch et al., 2023). Consequently, they may not see themselves as academically oriented, may have lower educational aspirations, and may make educational decisions (such as taking advanced courses and applying to selective colleges) that undermatch their abilities (Oyserman, 2013).

To make educators, students, and their families more aware when students are succeeding and to encourage them to keep striving, the Massachusetts Department of Elementary and Secondary Education (DESE) launched the Massachusetts Student Achievement Awards (MSAA) program in 2022. The initiative identified disadvantaged students performing at the “exceeding expectations” level on the state’s standardized tests (roughly the top 7%) in mathematics or English language arts (ELA) and sent them a certificate of achievement, a personalized letter from the Education Commissioner, informational resources, and a \$20 gift card.

Can a low-cost recognition nudge shift the achievement trajectory of these students? This study addresses this question using state longitudinal data and a difference-in-differences (DiD) design.<sup>1</sup> We find that students who received MSAA awards in grades 3–7 scored 0.05–0.07

---

<sup>1</sup> This report is a first look. A forthcoming study will use a regression discontinuity design, comparing disadvantaged students just above and below the 530 award threshold, to estimate MSAA’s effects on achievement,

standard deviations (SD) higher on mathematics tests the following year than they would have in the absence of the awards. The mathematics effects we find are larger for girls and for students who were not consistently among the state's top performers, suggesting that recognition nudges may be especially beneficial for students less aware of their academic achievement. By contrast, the MSAA effects on ELA achievement, while generally positive, are smaller and not consistently significant across model specifications.

We focus our main discussion on the mathematics findings and report the ELA results in Appendix A. Section 2 describes the MSAA program and its theory of change. Section 3 presents the data, sample, and empirical strategy. Section 4 reports the findings, and Section 5 concludes. In Appendix B, we focus on an even more selective group of high achievers—students who placed in the top 2% statewide in mathematics—and examine MSAA’s effect on these students’ subsequent mathematics achievement.

## **2. The Massachusetts Student Achievement Awards**

### ***2.1 Program Design and Implementation***

Under the MSAA program, awards were given to students with disabilities, English learners, and economically disadvantaged students who scored at the “exceeding expectations” level in mathematics or ELA (state standardized tests known as MCAS) or “progressing” in MCAS-Alt in Grades 3–8 and 10. Award recipients received a certificate signed by the Education Commissioner and state board chair, a personalized congratulatory letter signed by the Commissioner, a list of resources for supporting a high achieving student,<sup>2</sup> and a \$20 gift card.

---

attendance, and advanced course-taking. The RD estimates will offer stronger internal validity for students near the threshold, while the approach used here speaks to the broader population of high-achieving disadvantaged students.  
<sup>2</sup> These include tips on setting high expectations, promoting a growth mindset, celebrating learning, and seeking out challenges. A QR code was provided for a free subscription to The Busy Family’s Guide to School.

The total cost per award recipient was less than \$35, with the majority covering the gift card. Other costs included printing, postage, and project management.

The program sent awards to two cohorts of students who met all criteria. These include 12,982 students based on the Spring 2022 scores and 14,818 students based on the Spring 2023 scores. Awards were distributed by a contractor to the state, *EdNavigator*, in the winter of the subsequent school year through a mailing to each student’s school placement at that time. **Table 1** shows the breakdown by year, test type, and grade. Approximately one-third of the awards were based on MCAS-Alt, the alternate assessments for students with significant cognitive disabilities. Because our data lack scale scores for MCAS-Alt assessments, we exclude MCAS-Alt-based award recipients from the study. Among MCAS-based award recipients, 10th-grade students accounted for the largest share of award recipients, receiving 20% of the awards. **Figure 1** depicts the percentage of students receiving an award by test score (the higher score of the two subjects, centered at the award eligibility threshold). No student below the threshold (the vertical line at 0) received awards and nearly all disadvantaged students at or above the threshold received awards, indicating near perfect fidelity in program implementation.

**Table 2** reports selected sample statistics for award recipients and non-recipients. Consistent with the program design, all award recipients are in one or more of the three disadvantaged categories compared to 56% of other students. Economically disadvantaged students (78% among recipients *vs.* 45% among other students), students eligible for special education programs (45% *vs.* 21%), and Asian students (15% *vs.* 7%) are overrepresented amongst recipients. About 15% of the award recipients are “double” qualifiers, achieving “exceeding expectations” on both mathematics and ELA. About 48% of the award recipients qualified based on ELA only, and about 37% qualified based on mathematics only.

## **2.2     *Theory and Evidence of Nudge Initiatives***

The idea of nudges draws on long-standing insights about human decision-making. In 1955, Simon (1955) introduced the concept of bounded rationality, the idea that decision-making is constrained by information and cognitive limits that prevent people from making purely rational decisions. Given these constraints, Tversky and Kahneman (1974) argue that people rely on shortcuts or “judgmental heuristics” (including representativeness, availability, and adjustment and anchoring) that lead to systematic and predictable errors. Thaler and Sunstein (2008) suggest that small nudges like reminders and simplified processes can improve decision-making and guide behavior without restricting people’s freedom of choice.

Through recognition nudges, MSAA could potentially support more efficient decision-making by improving students’ awareness of their achievement and boosting self-efficacy. These cognitive and psychological benefits are thought to influence student motivation and expectations (Schunk & DiBenedetto, 2023; Urhahne & Wijnia, 2023).

Empirical evidence from randomized controlled trials (RCTs) shows the potential of nudges in education but also, more recently, their limits. For example, texting parents the number of days their child had missed school was found to decrease chronic absenteeism among K–5 students by 15% (Robinson et al., 2018). When parents were randomly selected to receive text messages about summer reading and literacy development techniques and resources, their third- and fourth-grade children improved reading comprehension by over 0.2 standard deviations (Kraft & Monti-Nussbaum, 2017). Ninth-graders with standardized test scores of “meets standards” or higher who were randomly selected to receive an informational nudge about their readiness for advanced coursework were more likely to plan advanced courses for later in high school. The nudges also increased participation in 10th grade advanced courses for English

learners—a group of students who may have less information about their readiness (Austin et al., 2026). The high school to college transition (including SAT/ACT test taking; financial aid application and renewal; and college selection, application and matriculation) is an active field of nudge experiments (e.g., Hoxby & Avery, 2013; Bettinger et al., 2012; Castleman & Page, 2015). While many studies have found positive effects of nudge interventions, some evidence suggests that nudges alone may not work (Bettinger et al., 2012) and that the nudge effect could vary by the message itself, the timing of the intervention, and the scale and the context of the intervention (Avery et al., 2021).

### **3. Data, Sample, and Empirical Strategy**

#### **3.1 Data and Sample**

Our study uses longitudinal administrative data from the Massachusetts Department of Elementary and Secondary Education (DESE), combining information from the Student Information Management System (SIMS) with MCAS assessment records. Data span the 2016–17 through 2023–24 school years except for 2019–20, when state testing was suspended due to the COVID-19 pandemic.

The main analytic sample for the mathematics analysis includes students in grades 3–7 who scored at the “exceeding expectations” level on MCAS mathematics in a pre-intervention year (spring 2017, 2018, or 2021) or a post-intervention year (spring 2022 or 2023). Similarly, the analytic sample for ELA includes students who exceeded expectations on MCAS ELA in the same grade and year range. Students in Grades 8 and 10 are excluded because there are no state standardized tests in subsequent grades. Students with significant cognitive disabilities who took the MCAS-Alt assessments are also excluded because scale scores are unavailable.

Disadvantaged students in the analytic sample constitute the treatment group, and non-disadvantaged students in this sample are the comparison group.

### 3.2 Empirical Strategy

Among students whose test scores exceeded expectations, disadvantaged students received the award in post-intervention years (2022, 2023), while non-disadvantaged students did not. Using pre-program cohorts (2017, 2018, 2021) as the baseline, we estimate the difference in subsequent year's scores between disadvantaged and non-disadvantaged students and examine how the difference changed from the pre- to the post-intervention years.

Specifically, we estimate a difference-in-difference (DiD) model of the following form using repeated cross-sectional data:

$$y_{ic} = \beta_0 + \beta_1 Post_c + \beta_2 T_i + \beta_3 Post_c * T_i + \sum_k \gamma_k * Covariates_{ic} + g + y + s + e_{ic}, (1)$$

A student  $i$  in cohort  $c$ 's outcome ( $y_{ic}$ , next year's score standardized by year and grade) is modeled as a function of whether the student is in a post-intervention cohort ( $Post_c = 1$ ), whether the student is eligible to receive an award ( $T_i = 1$ ), and the interaction of these two indicator variables. Equation (1) is estimated separately for mathematics and ELA. To account for compositional change of student population across cohorts, the model controls for student characteristics including race/ethnicity, gender, and indicators for disabilities, English learner eligibility, and economic disadvantage. The model also includes grade ( $g$ ), year ( $y$ ), and school ( $s$ ) fixed effects to account for unobserved heterogeneity across these units. The model is estimated using ordinary least squares with standard errors clustered at the school level.

The coefficient for the interaction term,  $\beta_3$ , represents the MSAA effect on high-achieving disadvantaged students, provided that the difference in achievement between the treatment and comparison groups would have remained unchanged in the post-intervention period in the absence of MSAA. One possible threat to this assumption is the existence of concurrent policies or interventions that also affected disadvantaged students' achievement in the

post-intervention period. In that case,  $\beta_3$  would capture both the MSAA effect and the effects of those concurrent policies. To address this concern, we estimate the same DiD model for students unaffected by MSAA: those scoring at the “meeting expectations” level, one tier below the award threshold, and therefore ineligible for awards. We then subtract the DiD estimate for the meeting expectations group from the DiD estimate for the exceeding expectations group in a triple-difference (DDD) framework, thereby netting out the potential influence of other concurrent policies.

Another threat to the causal interpretation of  $\beta_3$  is that the definition of “low income” changed starting in academic year 2021–22. The definition was expanded to include MassHealth (Medicaid) participants up to 185% of the federal poverty level, students identified by districts as homeless, and students verified by districts through a supplemental documentation process.<sup>3</sup> The percentage of disadvantaged students in Grades 3–7 increased from 50% to 57% between 2019–20 and 2021–22. If newly classified low-income students were higher achieving than students classified under the prior definition, disadvantaged students’ spring 2022 and later scores could rise mechanically because of the definitional change. Assuming this change affected the achievement comparison between disadvantaged and non-disadvantaged students similarly in both the “exceeding expectations” and the “meeting expectations” groups, the DDD design discussed earlier can difference out changes attributable to the low-income redefinition and isolates the MSAA effect.

## **4. Findings**

### **4.1 Descriptive Findings**

---

<sup>3</sup> See [Redefining Low-income Under the Student Opportunity Act \(SY 2021–22\) - Information Services/Data Collection](#) for more details.

**Table 3** shows substantial differences between the analytic sample—students who exceeded state expectations, regardless of disadvantaged status—and the broader student population. White and Asian students are overrepresented in the analytic sample, accounting for 65% and 19% of students, respectively, compared with 56% and 6% of the broader student population. In contrast, Black students (3% vs. 10%) and Hispanic students (7% vs. 24%) are underrepresented. Students who exceeded state expectations were consistently higher achieving than other students, scoring at least 1.4 SD higher on both prior and subsequent tests. The groups also differ sharply by disadvantaged status: 15% of the “exceeding expectations” group are disadvantaged, compared with 54% of other students. This underrepresentation at the upper end of the achievement distribution appears across all disadvantage categories, including special education status (3% vs. 22%), low-income status (12% vs. 42%), and English learner status (1% vs. 11%).

The underrepresentation of disadvantaged students is exacerbated by their declining share among high achievers across grades, dropping from 18% to 11% between Grades 3 and 7 in mathematics (**Figure 2**). The falling performance trajectory of disadvantaged students is also reflected in year-to-year test score changes. Although mathematics scores decline among all top achievers, possibly reflecting regression to the mean (**Figure 3**), they decline more for disadvantaged students (solid lines) than their non-disadvantaged peers (dashed lines). The gap in decline between the two groups appears narrower in the post-MSAA period (black lines) than in the pre-MSAA period (grey lines), suggesting a possible positive award effect.

#### **4.2 Regression Results**

DiD estimates for mathematics are reported in **Table 4**. Across all specifications, the MSAA effect on next year’s mathematics score ( $\text{Post} \times \text{T}$ ) is positive and statistically significant,

ranging from 0.06 to 0.08 student level SDs. The limited difference between the first two columns indicates that adding student-level covariates contributes little explanatory power beyond grade, year, and school fixed effects. Controlling for lagged mathematics test scores (Column 3) allows us to compare the outcomes among students with two years of similar baseline scores, though this specification limits the sample to students in Grades 4–7 who took MCAS in spring 2019, 2022, and 2023. The estimated MSAA effect (0.06 SD) is largely unchanged. The coefficient for the disadvantaged indicator (T) shows that disadvantaged students fell behind non-disadvantaged students in subsequent mathematics achievement by approximately 0.10 SD in the baseline period. Combined with the estimated MSAA effect, these numbers suggest that MSAA significantly slowed but may not have completely stemmed disadvantaged students’ achievement decline.

Our primary analytic sample includes three pre-intervention years: 2017, 2018, and 2021. Columns (4) and (5) assess whether the results are sensitive to the choice of baseline years. Column (4) excludes 2021 because it was the first post-COVID-19 testing year and the only year with MCAS participation below 90%, raising concerns about the representativeness of test takers. Column (5) excludes the pre-pandemic years, 2017 and 2018, out of the concern that pandemic-related learning loss may have fundamentally altered the achievement landscape, making pre-pandemic performance a less appropriate counterfactual for post-pandemic outcomes. Although neither specification is ideal, the estimated MSAA effects on mathematics achievement remain positive at 0.06 and 0.08 SD.

#### **4.3 *Triple-Difference Estimates***

We use triple difference models to examine the robustness of the estimated MSAA effect to the possible influence of other concurrent policies and the redefinition of “low income”. The

results are reported in **Table 5**. All models use the same fixed effects and student covariates as Table 4. The coefficient on the interaction between the post-intervention and disadvantaged group indicators ( $\text{Post} \times \text{T}$ ), now representing the change in relative mathematics achievement between disadvantaged and non-disadvantaged students in the “meeting expectations” group, is positive and statistically significant in all models. These estimates indicate that factors other than MSAA have modestly increased disadvantaged students’ mathematics achievement in the post-intervention period. Even so, gains are significantly larger among students who exceeded expectations. The DDD estimates ( $\text{Post} \times \text{T} \times \text{Z}$ ) range from 0.04 to 0.06 SDs across specifications, indicating net positive benefits of the MSAA program.

#### **4.4 Subgroup Results**

**Figure 4** presents the DDD estimates of the MSAA effect for various student subgroups. The effects were estimated with student covariates as well as grade (except for grade-specific groups), year, and school fixed effects. Student groups are listed on the y-axis. Dots represent point estimates, and whiskers represent the 95% confidence interval. The estimated effect for all students (0.05 SD, Table 5, Column 2) is reproduced at the top for reference.

The MSAA effect is slightly larger for girls (0.064) than for boys (0.045). One potential explanation for the larger effect is that girls may derive greater benefits from external validation of their mathematics abilities, as girls are widely documented to underestimate their mathematics competence relative to boys (e.g., Eccles et al., 1993, Correll, 2001). Even as the gender gap in mathematics achievement has narrowed in recent decades, a gap in mathematics self-confidence persists (Zander et al, 2020), and parents also continue to underestimate girls’ mathematics abilities (Brenøe & Lundberg, 2025).

The MSAA effect is positive across all racial/ethnic groups, though its magnitude varies: it is largest for Black students (0.186), followed by Asian (0.065), Hispanic (0.041), and White students (0.036). We also compare effects for students who exceeded expectations in two consecutive years against those who did so only once, reasoning that repeat top performers may already be more aware of their achievement, leaving less room for a recognition nudge to matter at the margin. Consistent with this, the estimated effect is smaller for repeat top performers than for one-time top performers (0.045 vs. 0.061), though the difference is not statistically significant.

The empirical literature on nudges suggests that their effects could vary by the timing of the intervention (Avery et al., 2021). Recognizing students in earlier grades may be more effective because it establishes high expectations before achievement patterns become more firmly established. Grade-specific estimates, however, do not appear to follow a clear pattern. Due to smaller sample sizes, only 3rd- and 5th-grade effects remain statistically significant and positive (0.057 and 0.068 SD, respectively). Fourth grade has the lowest point estimate (0.020), but it is not statistically different from other grade-specific effects.

Student mobility varies substantially across grades in our sample, largely because many fifth-grade students transition from elementary to middle school. Fifty-eight percent of 5th grade students attended a different school the following year, compared with 25% of 4th graders, 23% of 6th graders, 14% of 3rd graders, and 11% of 7th graders. Because recognition may operate differently when students remain in the same school than when they receive the award after changing schools, we estimate MSAA effects by school-switching status. The results provide little evidence that the award effect varies between movers and non-movers.

## **5. Discussion and Conclusion**

This study examines whether recognizing exceptional achievement can help sustain the performance of disadvantaged students who have demonstrated high academic potential. In Massachusetts, disadvantaged students are underrepresented among the state's highest-scoring students, and those who do reach elite levels are more likely than similarly high-performing non-disadvantaged peers to lose ground over time. The MSAA program sought to recognize these students through certificates, letters, informational resources, and a modest gift card. Using a DiD design and statewide longitudinal data, we find that the program increased next-year mathematics achievement by approximately 0.05–0.07 standard deviations among disadvantaged students who exceeded state expectations in grades 3–7. By contrast, we find little comparable evidence of an effect in ELA (Appendix A), despite similar patterns of underrepresentation and subsequent achievement decline among high-scoring disadvantaged students.

The mathematics findings are notable for their consistency. Estimated effects remain remarkably stable across specifications, comparison years, and modeling choices. The estimated effect persists after controlling for twice-lagged achievement and in a triple-difference framework that compares award recipients with students just below the eligibility threshold, reducing concerns that the results reflect the influence of other concurrent policies and the 2021–22 redefinition of low-income status. Although modest in absolute terms (Kraft, 2020), an effect of this magnitude is substantively important, equating to roughly one additional month of learning in elementary and secondary grades (Lipsey et al., 2012). Given that MSAA costs approximately \$35 per recipient and required little additional implementation infrastructure, it appears to be highly cost-effective.

The subgroup analyses provide additional suggestive evidence. Estimated effects are larger for girls than boys (0.064 vs. 0.045) and larger for students who exceeded state

expectations only once than for students who consistently performed at that level (0.061 vs. 0.045). One interpretation is that recognition is most valuable for students who are less likely to recognize their own exceptional achievement or whose achievement has not yet become an established part of their academic identity. While these differences are not statistically significant and cannot provide a direct test, the patterns are consistent with the hypothesis that recognition helps students by reducing information gaps about academic potential.

A couple of limitations warrant caution in interpreting these results. First, the sample of disadvantaged students who exceeded expectations on state standardized tests is relatively small, which limits statistical power for some subgroup analyses. Some of the grade and race/ethnicity estimates should be interpreted with appropriate uncertainty. Second, this study is limited to next year's test scores, precluding analysis of longer-term effects on course-taking or high school outcomes. Future research should examine these longer-term effects. Richer data on when families received and engaged with the materials would help distinguish between the recognition, information, and financial incentive components of the award.

This report provides a first look at MSAA's effects. In a forthcoming study, we will use a regression discontinuity (RD) design to compare disadvantaged students who scored just above the award eligibility threshold (a scale score of 530 in mathematics or ELA) with those who scored just below it. Because scoring at or above 530 also earns students the "exceeding expectations" label, the RD analysis will net out the discontinuity at the same threshold in pre-program cohorts. That study will also extend the outcomes beyond achievement to include attendance and advanced course-taking. The RD design offers stronger internal validity, but its estimates apply only to students near the threshold. The difference-in-differences approach used

here covers the full population of high-achieving disadvantaged students and will serve as a complement to the RD estimates.

Despite these limitations, the findings point to an overlooked opportunity in education policy. States devote substantial attention to identifying students who are struggling academically but comparatively little attention to identifying and encouraging disadvantaged students who demonstrate exceptional talent. The evidence presented here suggests that a simple, inexpensive, and easily scalable recognition program can modestly improve subsequent mathematics achievement among this small but critical group of students, with the potential to increase diversity in the ranks of top mathematics students. Whether similar programs would produce comparable results in other states—and why the effects appear in mathematics but not in ELA—are important questions for future research.

## References

- Anderson, K. P. (2024). Biased opportunities: The role of implicit and explicit bias in advanced placement and dual enrollment. *Educational Evaluation and Policy Analysis*, 46(4), 735–763.
- Austin, M., Farmer, M., Kruse, M., Mbekeani, P. P., Mitrano, M., & Nathenson, R. (2026). The effects of an automatic notification tool to increase participation in advanced high school courses: Results from a large-scale cluster randomized controlled trial. (EdWorkingPaper: 26-1427). Annenberg Institute at Brown University. <https://doi.org/10.26300/5s80-9t41>
- Avery, C., Castleman, B. L., Hurwitz, M., Long, B. T., & Page, L. C. (2021). Digital messaging to improve college enrollment and success. *Economics of Education Review*, 84, 102170. <https://doi.org/10.1016/j.econedurev.2021.102170>
- Batruch, A., Geven, S., Kessenich, E., & van de Werfhorst, H. G. (2023). Are tracking recommendations biased? A review of teachers' role in the creation of inequalities in tracking decisions. *Teaching and Teacher Education*, 123, 103985.
- Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics*, 127(3), 1205–1242. <https://doi.org/10.1093/qje/qjs017>
- Brenøe, A. A., & Lundberg, S. (2025). Peers, parents, and self-perceptions: The gender gap in mathematics self-assessment. *Journal of Population Economics*, 38, 21. <https://doi.org/10.1007/s00148-025-01087-2>
- Carina Initiatives. (n.d.). About us. Retrieved June 9, 2026, from <https://www.carina.fund/about-us>
- Castleman, B. L., & Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115, 144–160. <https://doi.org/10.1016/j.jebo.2014.12.008>
- Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, 106(6), 1691–1730.
- Cunha, F. (2009). Crossing the finish line: A review article. *Journal of Human Capital*, 3(4), 354–378.
- Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64(3), 830–840.

- Goldhaber, D., Wolff, M., & Daly, T. (2021). Assessing the accuracy of elementary school test scores as predictors of students' high school outcomes (CALDER Working Paper No. 235-0821-2). American Institutes for Research.
- Hoxby, C. M., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, Spring, 1–65.
- Hoxby, C. M., & Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review*, 105(5), 514–517.  
<https://doi.org/10.1257/aer.p20151027>
- Koedel, C., & Betts, J. R. (2010). Value added to what? How a ceiling in the testing instrument influences value-added estimation. *Education Finance and Policy*, 5(1), 54–81.  
<https://doi.org/10.1162/edfp.2009.5.1.5104>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Kraft, M. A., & Monti-Nussbaum, M. (2017). Can schools enable parents to prevent summer learning loss? A text messaging field experiment to promote literacy skills. *The ANNALS of the American Academy of Political and Social Science*, 674(1), 85–112.  
<https://doi.org/10.1177/0002716217732009>
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms (NCSE 2013-3000). National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.  
[https://nces.ed.gov/sites/default/files/migrated/nces\\_pubs/ncser/pubs/20133000/pdf/20133000.pdf](https://nces.ed.gov/sites/default/files/migrated/nces_pubs/ncser/pubs/20133000/pdf/20133000.pdf)
- Oyserman, D. (2013). Not just any path: Implications of identity-based motivation for disparities in school outcomes. *Economics of Education Review*, 33, 179–190.  
<https://doi.org/10.1016/j.econedurev.2012.09.002>
- Page, L. C., & Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, 51, 4–22.  
<https://doi.org/10.1016/j.econedurev.2016.02.009>
- Robinson, C. D., Lee, M. G., Dearing, E., & Rogers, T. (2018). Reducing student absenteeism in the early grades by targeting parental beliefs. *American Educational Research Journal*, 55(6), 1163–1192. <https://doi.org/10.3102/0002831218772274>
- Schunk, D. H. & DiBenedetto, M. K. (2023). Learning from a social cognitive theory perspective. In R. J. Tierney, F. Rizvi, & K. Ercikan (Eds.), *International Encyclopedia of Education* (4th ed.) (pp. 22-35). Elsevier.

- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Urhahne, D., & Wijnia, L. (2023). Theories of motivation in education: An integrative framework. *Educational Psychology Review*, 35(2), 45. <https://doi.org/10.1007/s10648-023-09767-9>
- Zander, L., Höhne, E., Harms, S., Pfof, M., & Hornsey, M. J. (2020). When grades are high but self-efficacy is low: Unpacking the confidence gap between girls and boys in mathematics. *Frontiers in Psychology*, 11, 552355. <https://doi.org/10.3389/fpsyg.2020.552355>

## Tables and Figures

**Table 1. Number of Student Achievement Award Recipients, by Cohort, Grade, and Assessment Type**

| Grade | 2022          |                   |        | 2023          |                   |        |
|-------|---------------|-------------------|--------|---------------|-------------------|--------|
|       | MCAS<br>award | MCAS-Alt<br>award | Total  | MCAS<br>award | MCAS-Alt<br>award | Total  |
| 3     | 1,295         | 719               | 2,014  | 1,421         | 674               | 2,095  |
| 4     | 851           | 691               | 1,542  | 1,259         | 732               | 1,991  |
| 5     | 824           | 629               | 1,453  | 925           | 689               | 1,614  |
| 6     | 1,162         | 628               | 1,790  | 1,288         | 625               | 1,913  |
| 7     | 1,092         | 611               | 1,703  | 1,219         | 597               | 1,816  |
| 8     | 1,319         | 610               | 1,929  | 1,654         | 597               | 2,251  |
| 10    | 1,887         | 606               | 2,493  | 2,546         | 592               | 3,138  |
| Total | 8,430         | 4,494             | 12,924 | 10,312        | 4,506             | 14,818 |

Source: EdNavigator.

Note: This table excludes 58 2022 award recipients who do not have MCAS or MCAS-Alt records.

***Table 2. Comparison of Student Characteristics Between Award Recipients and Non-Recipients, 2022 and 2023 Cohorts***

| Characteristic             | Award recipients | Non- recipients |
|----------------------------|------------------|-----------------|
| Disadvantaged              | 1.00             | 0.56            |
| English learner            | 0.09             | 0.13            |
| Economically disadvantaged | 0.78             | 0.45            |
| Special education          | 0.45             | 0.21            |
| Male                       | 0.56             | 0.52            |
| Asian                      | 0.15             | 0.07            |
| Black                      | 0.11             | 0.10            |
| Hispanic                   | 0.23             | 0.25            |
| White                      | 0.45             | 0.54            |
| Observations               | 27,801           | 1,941,026       |

*Note:* The variable “disadvantaged” is a constructed indicator variable that equals 1 if a student is an English learner, economically disadvantaged, or in a special education program and 0 otherwise.

**Table 3. Comparison of Student Characteristics Between Students Exceeding Expectations in Mathematics or ELA and Other Students**

| Student characteristic            | Exceeding expectations | Other students  |
|-----------------------------------|------------------------|-----------------|
| Next year's math score (SD units) | 1.33<br>(0.71)         | -0.14<br>(0.89) |
| Next year's ELA score (SD units)  | 1.25<br>(0.66)         | -0.13<br>(0.91) |
| Prior math score (SD units)       | 1.30<br>(0.70)         | -0.16<br>(0.91) |
| Prior ELA score (SD units)        | 1.22<br>(0.71)         | -0.15<br>(0.92) |
| Disadvantaged                     | 0.15<br>(0.36)         | 0.54<br>(0.50)  |
| English learner                   | 0.01<br>(0.10)         | 0.11<br>(0.31)  |
| Low income                        | 0.12<br>(0.32)         | 0.42<br>(0.49)  |
| Special education                 | 0.03<br>(0.17)         | 0.22<br>(0.41)  |
| Male                              | 0.45<br>(0.50)         | 0.52<br>(0.50)  |
| Asian                             | 0.19<br>(0.39)         | 0.06<br>(0.23)  |
| Black                             | 0.03<br>(0.17)         | 0.10<br>(0.30)  |
| Hispanic                          | 0.07<br>(0.26)         | 0.24<br>(0.43)  |
| White                             | 0.65<br>(0.48)         | 0.56<br>(0.50)  |
| Observations                      | 187,495                | 1,474,924       |

*Note:* Standard deviation in parentheses. Sample restricted to Grades 3–7, standard MCAS test takers in spring 2017, 2018, 2021, 2022, and 2023 with a subsequent-year MCAS record. Lagged scores further restricted to Grades 4–7 test takers in spring 2018, 2022, and 2023. “Disadvantaged” equals 1 if student is an English learner, economically disadvantaged, or in a special education program.

**Table 4. Difference-in-Differences (DiD) Estimates: The MSAA Effect on Next Year's Mathematics Scores**

|                      | (1)                | (2)                | (3)                | (4)                | (5)                |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                      | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023) | 0.12***<br>(0.01)  | 0.10***<br>(0.01)  | 0.03***<br>(0.01)  | 0.11***<br>(0.01)  | -0.05***<br>(0.01) |
| Disadvantaged (T)    | -0.13***<br>(0.01) | -0.08***<br>(0.03) | -0.07*<br>(0.04)   | -0.08***<br>(0.03) | -0.11***<br>(0.04) |
| Post × T (DiD)       | 0.07***<br>(0.01)  | 0.07***<br>(0.01)  | 0.06***<br>(0.02)  | 0.06***<br>(0.01)  | 0.08***<br>(0.02)  |
| Lagged score         |                    |                    | 0.38***<br>(0.01)  |                    |                    |
| English learner      |                    | -0.16***<br>(0.02) | -0.09*<br>(0.04)   | -0.15***<br>(0.02) | -0.17***<br>(0.03) |
| Low income           |                    | -0.05**<br>(0.02)  | -0.01<br>(0.04)    | -0.05*<br>(0.03)   | -0.04<br>(0.03)    |
| Special education    |                    | -0.04<br>(0.02)    | 0.03<br>(0.03)     | -0.03<br>(0.03)    | -0.01<br>(0.03)    |
| Male                 |                    | 0.04***<br>(0.00)  | 0.01*<br>(0.00)    | 0.04***<br>(0.00)  | 0.04***<br>(0.01)  |
| Asian                |                    | 0.15***<br>(0.01)  | 0.10***<br>(0.01)  | 0.15***<br>(0.01)  | 0.14***<br>(0.01)  |
| Black                |                    | -0.15***<br>(0.02) | -0.07***<br>(0.02) | -0.14***<br>(0.02) | -0.14***<br>(0.02) |
| Hispanic             |                    | -0.14***<br>(0.01) | -0.06***<br>(0.02) | -0.13***<br>(0.01) | -0.15***<br>(0.02) |
| White                |                    | -0.11***<br>(0.01) | -0.07***<br>(0.01) | -0.10***<br>(0.01) | -0.11***<br>(0.01) |
| N                    | 104,599            | 104,599            | 47,540             | 89,443             | 55,423             |
| R-squared            | 0.14               | 0.17               | 0.30               | 0.17               | 0.18               |

*Note:* All models include grade, year, and school fixed effects. Sample restricted to students who exceeded expectations in math in grades 3–7. Standard errors clustered at the school level in parentheses.  
 \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

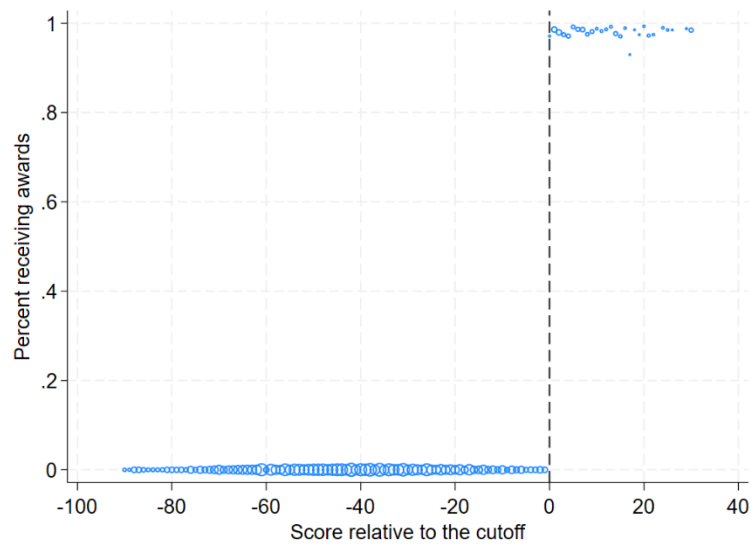
**Table 5. Triple-Difference (DDD) Estimates: The MSAA Effect on Next Year's Mathematics Scores**

|                                  | (1)                | (2)                | (3)                | (4)                | (5)                |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                                  | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023)             | 0.09***<br>(0.00)  | 0.09***<br>(0.00)  | 0.03***<br>(0.01)  | 0.09***<br>(0.00)  | -0.10***<br>(0.00) |
| Disadvantaged (T)                | -0.21***<br>(0.00) | -0.03***<br>(0.01) | -0.03***<br>(0.01) | -0.04***<br>(0.01) | -0.07***<br>(0.01) |
| Exceed expectations (Z)          | 0.84***<br>(0.00)  | 0.80***<br>(0.00)  | 0.49***<br>(0.01)  | 0.82***<br>(0.00)  | 0.75***<br>(0.01)  |
| Post $\times$ T                  | 0.02***<br>(0.00)  | 0.02***<br>(0.00)  | 0.03***<br>(0.01)  | 0.02***<br>(0.00)  | 0.05***<br>(0.01)  |
| Post $\times$ Z                  | 0.02***<br>(0.01)  | 0.02***<br>(0.01)  | -0.01<br>(0.01)    | 0.00<br>(0.01)     | 0.06***<br>(0.01)  |
| T $\times$ Z                     | 0.05***<br>(0.01)  | 0.04***<br>(0.01)  | 0.03*<br>(0.01)    | 0.04***<br>(0.01)  | 0.05***<br>(0.02)  |
| Post $\times$ T $\times$ Z (DDD) | 0.05***<br>(0.01)  | 0.05***<br>(0.01)  | 0.05***<br>(0.02)  | 0.06***<br>(0.01)  | 0.04**<br>(0.02)   |
| N                                | 691,726            | 691,726            | 323,250            | 586,060            | 363,551            |
| R-squared                        | 0.33               | 0.35               | 0.53               | 0.35               | 0.35               |

*Note:* Z = 1 for students exceeding expectations; Z = 0 for students meeting expectations who are therefore not eligible to receive MSAA. All models control for the same student characteristics and fixed effects as specified in model (2) in Tables 4. Standard errors clustered at the school level in parentheses.

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

**Figure 1. Treatment Assignment Compliance: 2022 and 2023 MCAS Cohorts**



*Note:* Test score on the x-axis is the higher score between ELA and mathematics, centered at the cutoff (530) at or above which a student is deemed “exceeding expectations”. The size of the circle is proportional to the number of students who scored at a given point.

**Figure 2. Share Disadvantaged Among Students Exceeding Expectations in Mathematics, by Grade**

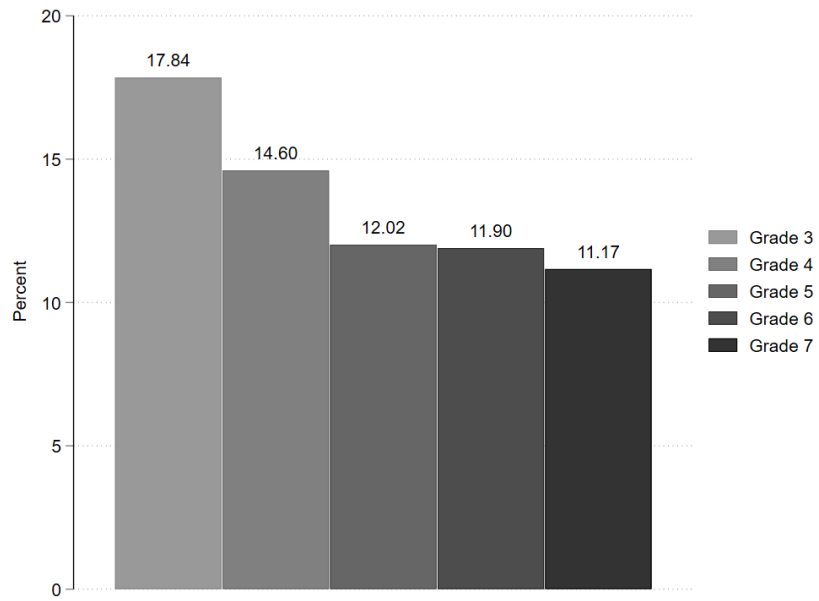


**Figure 3. Achievement Trajectory Among Students Who Exceeded Expectations in Mathematics in Period T, by Disadvantaged Status and Pre-Post Period**

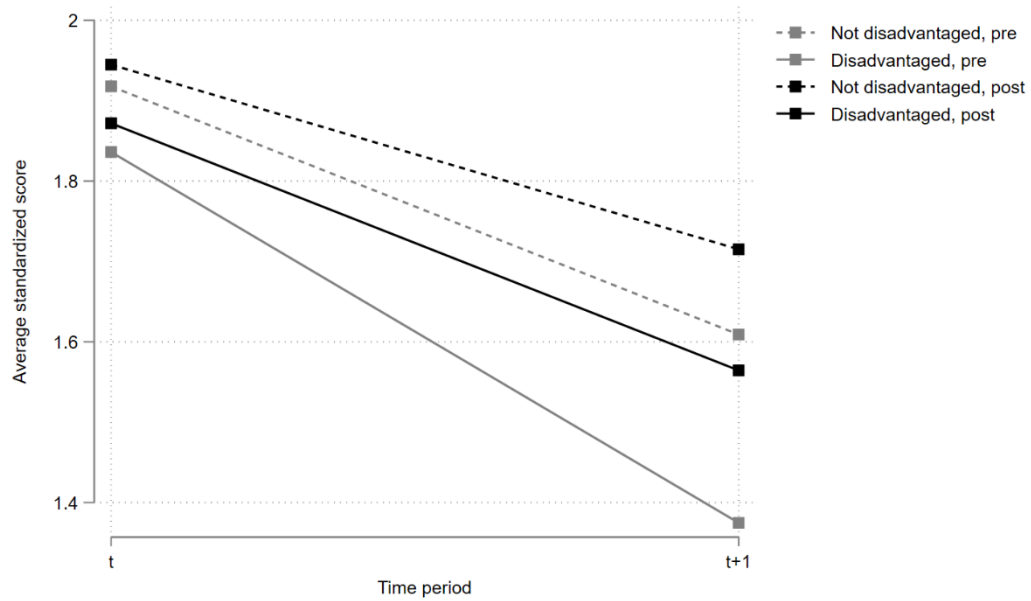
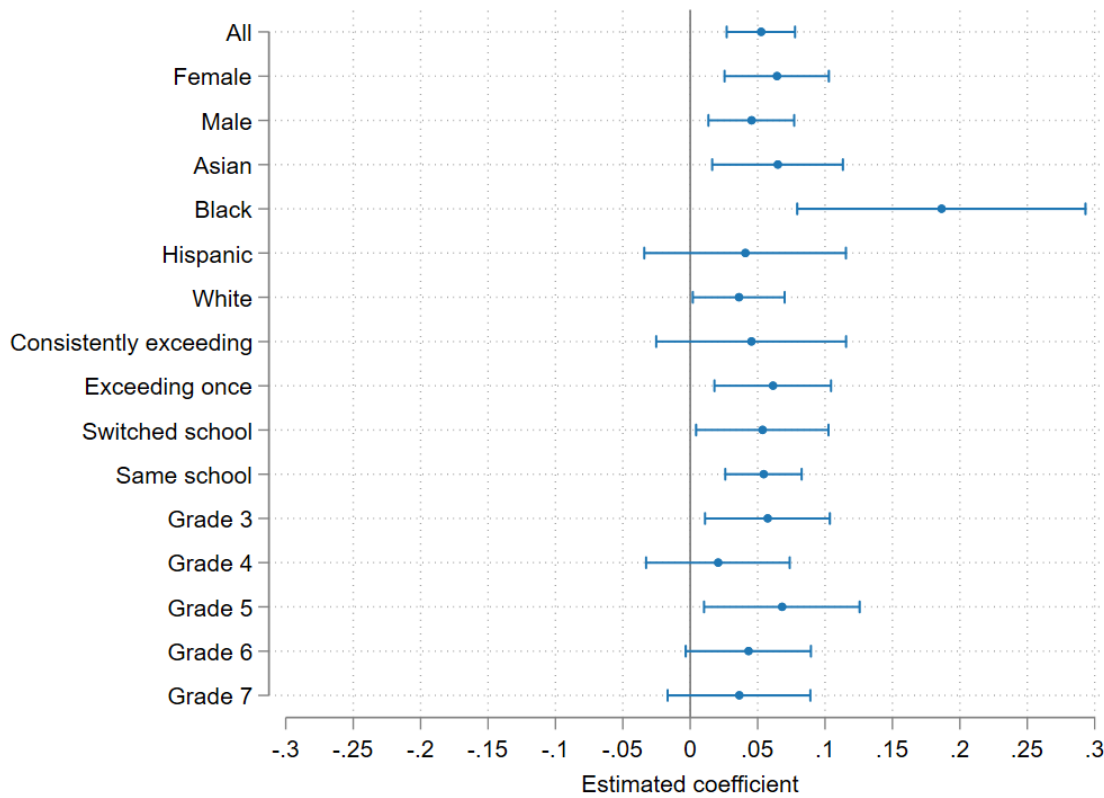


**Figure 4. DDD Estimates of the MSAA Effect on Next Year's Mathematics Achievement, by Student Group**



*Note:* Dots are the DDD estimates of the nudge effect in student standard deviation units, and whiskers represent the 95% confidence interval. The effects were estimated with student covariates as well as grade (except for grade-specific groups), year, and school fixed effects. Standard errors were clustered at the school level.

## Appendix A. ELA Results

This appendix reports a parallel set of results for ELA. Similar to mathematics, the share of disadvantaged students among those who exceeded state expectations in ELA declines from 17% in Grade 3 to 13% in Grade 7 (**Figure A1**), and ELA test scores tend to drop more year-to-year among disadvantaged high-performing students than among non-disadvantaged peers (**Figure A2**).

The estimated DiD effects of MSAA on ELA achievement are statistically significant across all model specifications and baseline choices, ranging from 0.03 to 0.06 SD (**Table A1**). However, DDD estimates indicate that disadvantaged students performing at the “meeting expectations” level experienced similar gains post-intervention (**Table A2**,  $\text{Post} \times \text{T}$ ), suggesting a null net effect of MSAA on ELA achievement. DDD estimates of the MSAA effect for various student subgroups are presented in **Figure A3**, and none of the effects are statistically significant.

**Table A1. Difference-in-Differences Estimates: Effect of MSAA on Next Year's ELA Scores**

|                      | (1)                | (2)                | (3)                | (4)                | (5)                |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                      | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023) | 0.11***<br>(0.01)  | 0.11***<br>(0.01)  | 0.11***<br>(0.01)  | 0.11***<br>(0.01)  | 0.15***<br>(0.01)  |
| Disadvantaged (T)    | -0.18***<br>(0.01) | -0.02<br>(0.03)    | -0.06<br>(0.04)    | -0.08**<br>(0.03)  | -0.02<br>(0.04)    |
| Post × T (DiD)       | 0.06***<br>(0.01)  | 0.05***<br>(0.01)  | 0.03*<br>(0.02)    | 0.06***<br>(0.01)  | 0.05***<br>(0.01)  |
| Lagged score         |                    |                    | 0.29***<br>(0.00)  |                    |                    |
| English learner      |                    | -0.24***<br>(0.04) | -0.06<br>(0.07)    | -0.18***<br>(0.04) | -0.24***<br>(0.05) |
| Low income           |                    | -0.12***<br>(0.03) | -0.02<br>(0.04)    | -0.07**<br>(0.03)  | -0.13***<br>(0.04) |
| Special education    |                    | -0.20***<br>(0.03) | -0.06<br>(0.04)    | -0.13***<br>(0.03) | -0.19***<br>(0.04) |
| Male                 |                    | -0.14***<br>(0.00) | -0.09***<br>(0.01) | -0.13***<br>(0.00) | -0.14***<br>(0.01) |
| Asian                |                    | 0.06***<br>(0.01)  | 0.04***<br>(0.01)  | 0.06***<br>(0.01)  | 0.04***<br>(0.01)  |
| Black                |                    | -0.15***<br>(0.02) | -0.12***<br>(0.02) | -0.14***<br>(0.02) | -0.16***<br>(0.02) |
| Hispanic             |                    | -0.16***<br>(0.01) | -0.10***<br>(0.01) | -0.15***<br>(0.01) | -0.16***<br>(0.02) |
| White                |                    | -0.08***<br>(0.01) | -0.06***<br>(0.01) | -0.07***<br>(0.01) | -0.10***<br>(0.01) |
| N                    | 118,483            | 118,483            | 52,992             | 92,494             | 65,340             |
| R-squared            | 0.09               | 0.11               | 0.23               | 0.11               | 0.13               |

*Note:* All models include grade, year, and school fixed effects. Sample restricted to students who exceeded expectations in ELA in grades 3–7. Standard errors clustered at the school level in parentheses.

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

**Table A2. Triple-Difference (DDD) Estimates: Effect of MSAA on Next Year's ELA Scores**

|                         | (1)                | (2)                | (3)                | (4)                | (5)                |
|-------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                         | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023)    | 0.12***<br>(0.01)  | 0.12***<br>(0.01)  | 0.10***<br>(0.01)  | 0.12***<br>(0.01)  | 0.11***<br>(0.00)  |
| Disadvantaged (T)       | -0.26***<br>(0.00) | -0.06***<br>(0.01) | -0.03***<br>(0.01) | -0.07***<br>(0.01) | -0.09***<br>(0.01) |
| Exceed expectations (Z) | 0.66***<br>(0.00)  | 0.63***<br>(0.00)  | 0.41***<br>(0.01)  | 0.65***<br>(0.00)  | 0.60***<br>(0.01)  |
| Post × T                | 0.05***<br>(0.00)  | 0.04***<br>(0.00)  | 0.01**<br>(0.01)   | 0.04***<br>(0.00)  | 0.06***<br>(0.01)  |
| Post × Z                | -0.01*<br>(0.01)   | -0.01<br>(0.01)    | -0.00<br>(0.01)    | -0.03***<br>(0.01) | 0.03***<br>(0.01)  |
| T × Z                   | 0.07***<br>(0.01)  | 0.05***<br>(0.01)  | 0.05***<br>(0.01)  | 0.05***<br>(0.01)  | 0.08***<br>(0.01)  |
| Post × T × Z (DDD)      | 0.01<br>(0.01)     | 0.01<br>(0.01)     | 0.01<br>(0.02)     | 0.02<br>(0.01)     | -0.01<br>(0.02)    |
| N                       | 755,420            | 755,420            | 337,364            | 607,357            | 412,181            |
| R-squared               | 0.23               | 0.25               | 0.39               | 0.26               | 0.26               |

*Note:* Z = 1 for students exceeding expectations; Z = 0 for students meeting expectations who are therefore not eligible to receive MSAA. All models control for the same student characteristics and fixed effects as specified in model (2) in Tables A1. Standard errors clustered at the school level in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

**Figure A1. Share Disadvantaged Among Students Exceeding Expectations in ELA, by Grade**

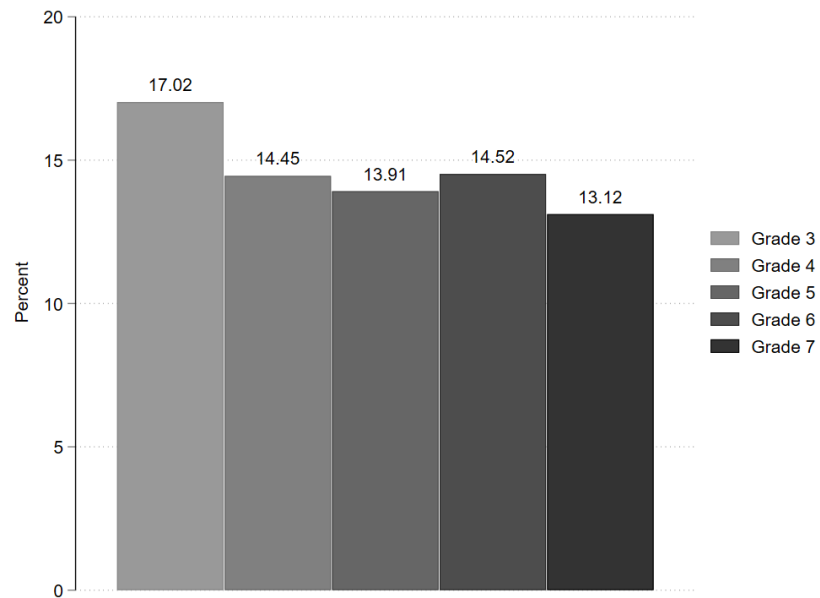


**Figure A2. Achievement Trajectory Among Students Who Exceeded Expectations in ELA in Period T, by Disadvantaged Status and Pre-Post Period**

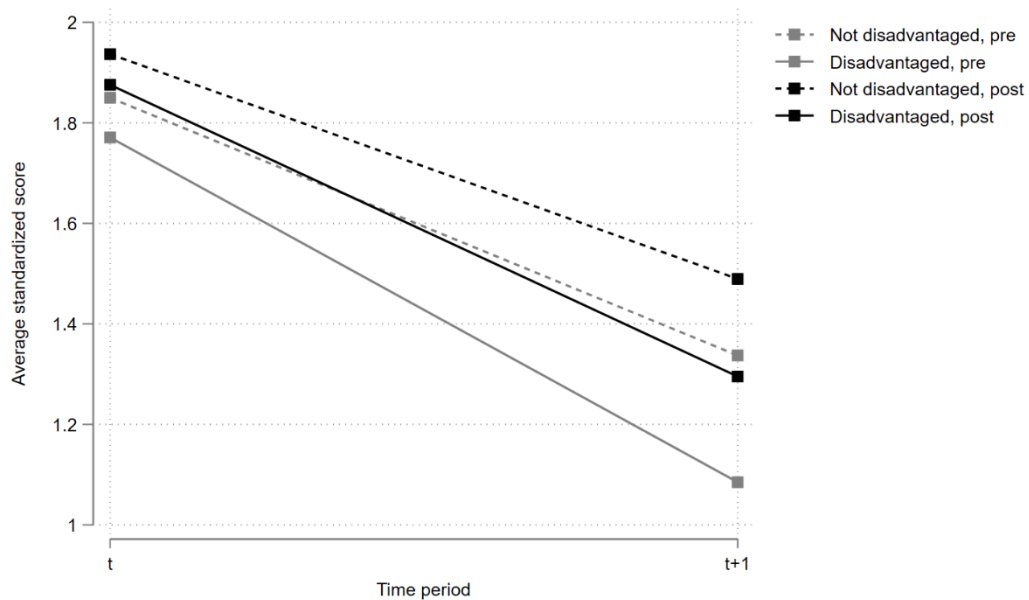
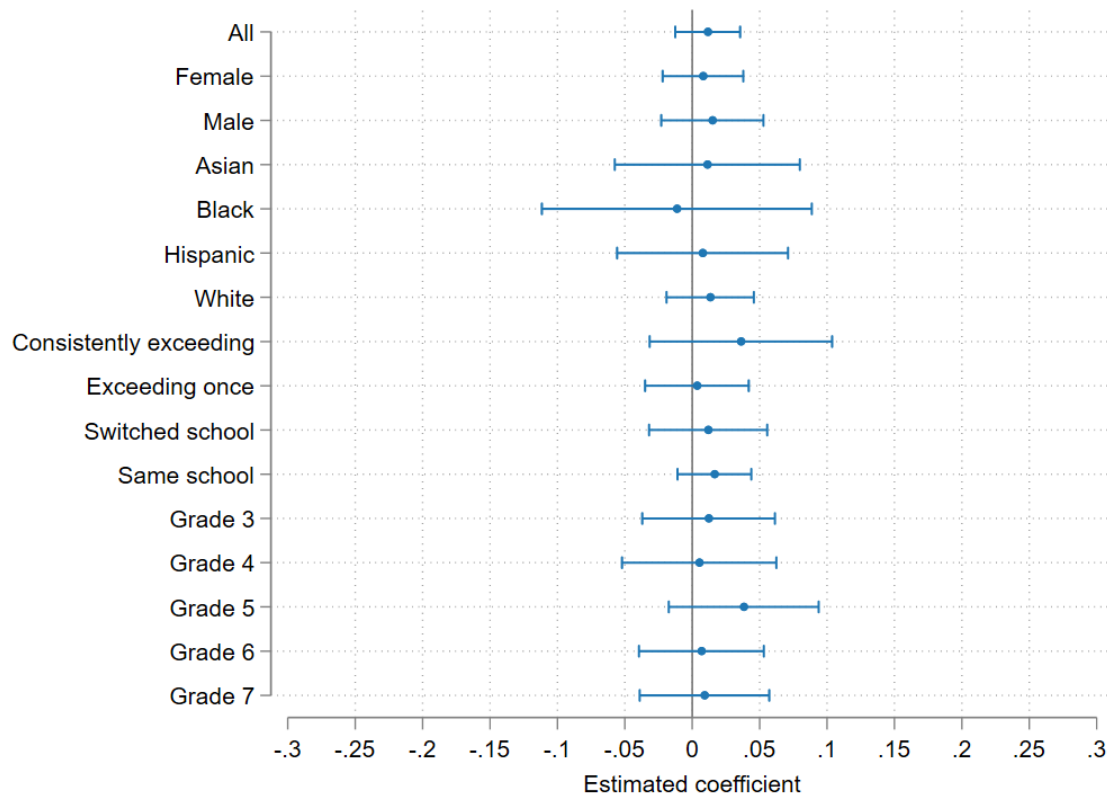


**Figure A3. DDD Estimates of the MSAA Effect on Next Year's ELA Achievement, by Student Group**



*Note:* Dots are the DDD estimates of the nudge effect in student standard deviation units, and whiskers represent the 95% confidence interval. The effects were estimated with student covariates as well as grade (except for grade-specific groups), year, and school fixed effects. Standard errors were clustered at the school level.

## Appendix B. Results for Top 2% Mathematics Achievers

Although children who demonstrate extraordinary mathematics abilities make disproportionately large contributions to innovation and STEM fields later, much of this talent has been neglected, especially among children from disadvantaged backgrounds (Carina Initiatives, n.d.). In this appendix, we focus on a subset of MSAA-eligible students who scored in the top 2% of the statewide mathematics distribution. Using the same data and methodology described in the main report, we find that students who received MSAA awards in grades 3–7 scored approximately 0.07 standard deviations (SD) higher on mathematics tests the following year than they would have absent the awards. The effect is larger for girls and for students not consistently in the top 2%, suggesting that recognition nudges may be especially beneficial for students less aware of their academic achievement.

A unique concern with focusing on top-performing students is the potential for a test score ceiling effect. This occurs when a test has limited differentiating power at the upper end, resulting in many students achieving the maximum or near-maximum score. The clustering of scores at the top would compress the variance of the dependent variable and attenuate the estimated coefficients towards 0. We examine the possibility of ceiling effects in MCAS mathematics in two ways. Panel A in **Figure B1** is a dot plot of the standardized MCAS mathematics scores of the top 2% students. In this very selective group, the scores spread out across values (the y-axis) rather than bunching at the top value. Another way to detect ceiling effects is to examine the skewness of the score distribution (Panel B). Crowding of high scores would result in a negative skewness. In an example of likely ceiling effects, Koedel and Betts (2010) found that Texas’s standardized tests at the time of the study had a skewness of -1.60 to -2.08, indicating that student scores were heavily concentrated at the upper end of the performance distribution, with a long tail extending toward lower scores. That is not the case

with MCAS mathematics tests, which have a very close to normal distribution with a skewness of -0.05.

### ***B.1 Descriptive Findings***

**Table B1** shows substantial differences between the top 2% analytic sample and the broader student population. In terms of racial composition, Asian students are overrepresented in the top 2%, making up 38% of top-2% students compared with 7% of non-top-2% students. By contrast, Black students (2% vs. 9%) and Hispanic students (4% vs. 24%) are underrepresented, while White students are slightly underrepresented (49% vs. 55%).

Top 2% students scored, on average, about 1.9 SD higher than other students on both prior and subsequent mathematics tests. The groups also differ sharply by disadvantaged status: 10% of top 2% students are disadvantaged, compared with 51% of non-top-2% students. This underrepresentation at the upper end of the achievement distribution appears across all disadvantage categories, including special education status (3% vs. 20%), low-income status (7% vs. 39%), and English learner status (1% vs. 10%). The share of disadvantaged students among top 2% performers falls as students progress through the grades, dropping from 14% to just over 8% between Grades 3 and 7 (**Figure B2**). The falling performance trajectory of disadvantaged students is also reflected in the steeper decline in test scores year-to-year (**Figure B3**, solid lines) relative to the score change among non-disadvantaged students (dashed lines). The gap in decline between disadvantaged and non-disadvantaged students, however, is narrower in the post-MSAA period, suggesting a possible positive award effect.

### ***B.2 Regression Results***

Our primary DiD estimates are reported in **Table B2**. The MSAA effect (Post x T) is positive and generally statistically significant with effect sizes ranging from 0.05 to 0.10 student

level SDs. Compared to a model that includes grade, year, and school fixed effects only (Column 1), adding student-level covariates (Column 2) results in no change in the estimated MSAA effect (0.07 SD). Adding twice-lagged mathematics test score increases the estimate to 0.09 SD.<sup>4</sup> Because the achievement gap between disadvantaged and non-disadvantaged students is more than 0.10 SD in the pre-intervention period (the coefficient on the disadvantaged indicator), MSAA significantly narrowed but may not have eliminated the achievement deficit.

Columns (4) and (5) assess whether the results are sensitive to the choice of baseline years for reasons discussed in the main report. The estimated MSAA effect is 0.05 when the first post-COVID year (2021) is excluded and 0.08 when the pre-COVID years are excluded. These estimates are within the range of the estimated MSAA effect using all baseline years as the pre-intervention period.

Finally, DDD estimates are nearly identical to the DiD estimates (**Table B3**). The coefficient for  $\text{Post} \times \text{T} \times \text{Z}$  ranges from 0.05 to 0.09 SD, suggesting that effects reported earlier are not driven by the influence of other concurrent policies or the redefinition of “low income”.

### ***B.3 Subgroup Results***

The MSAA effect was also estimated by various student subgroups using the DDD model (**Figure B4**). Most of the estimates are not statistically significant due to small sample sizes, but the patterns are similar to those reported in the main analysis. Specifically, the MSAA effect is slightly larger for girls than for boys (0.064 vs. 0.049 SD). It is also larger for students who were in the top 2% only once than for students in the top 2% twice (0.088 vs. -0.078). Black students in the top 2% benefited from the program the most with an estimated effect of 0.40 SD. What deviates slightly from the main results is the effect on school switchers. Among the top 2%

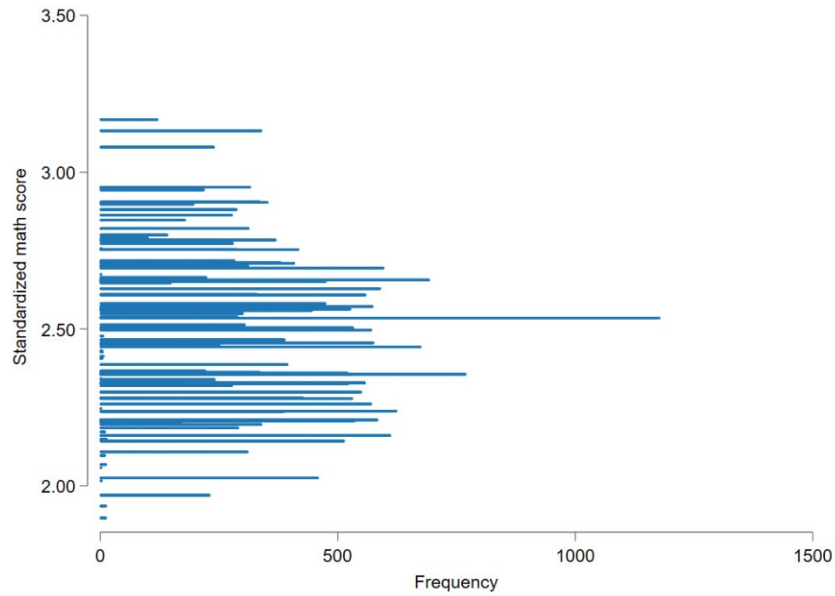
---

<sup>4</sup> This increase is unrelated to the changing sample. The MSAA effect estimated using the same specification as Column 2 with the smaller sample is 0.07 SD, identical to the full sample estimate.

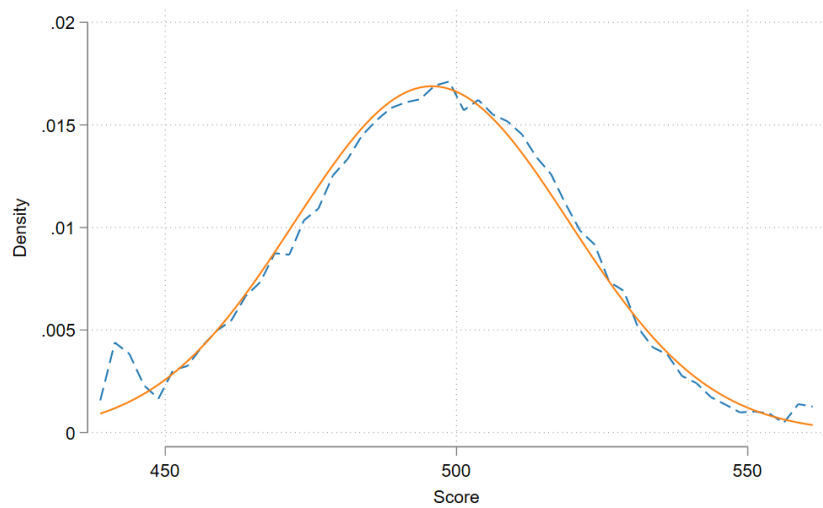
students, the MSAA effect for students who stayed in the same school is 0.084 and statistically significant, whereas the effect for students who moved to a different school is -0.049 (not statistically significant).

**Figure B1. MCAS Math Score Distribution, Spring 2022 and 2023**

(A) Dot plot of standardized math scores of the top 2%



(B) Kernel density estimate of math scale scores



— MCAS math score — Normal curve

Skewness = -.046

**Figure B2. Share Disadvantaged Among Top 2%, by Grade**

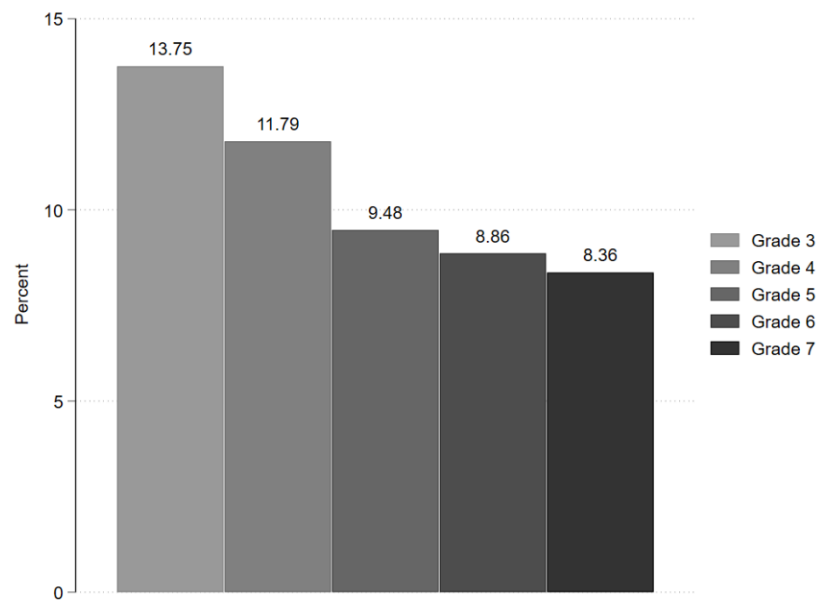


**Figure B3. Math Achievement Trajectory**

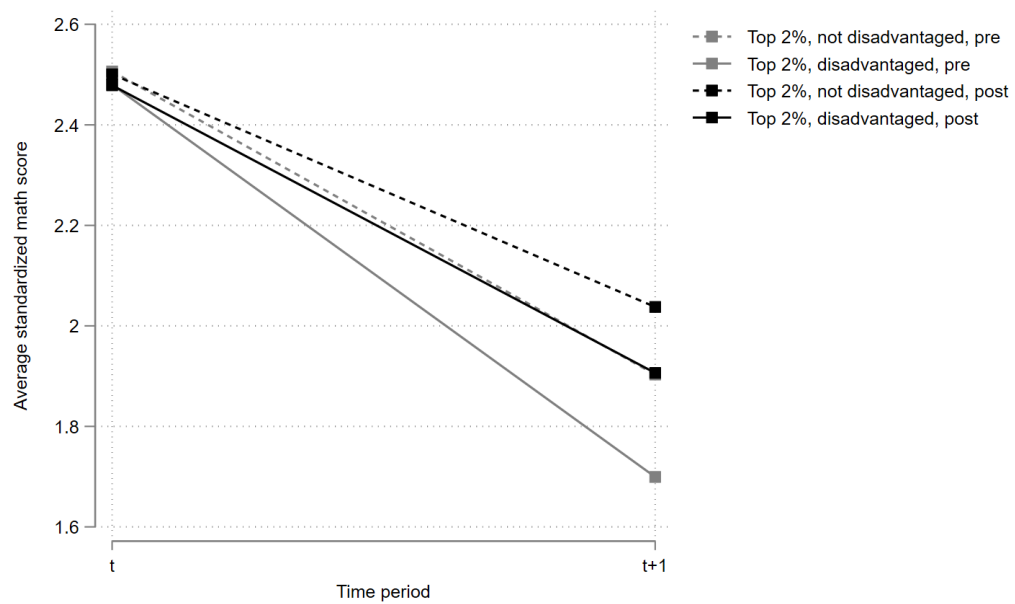
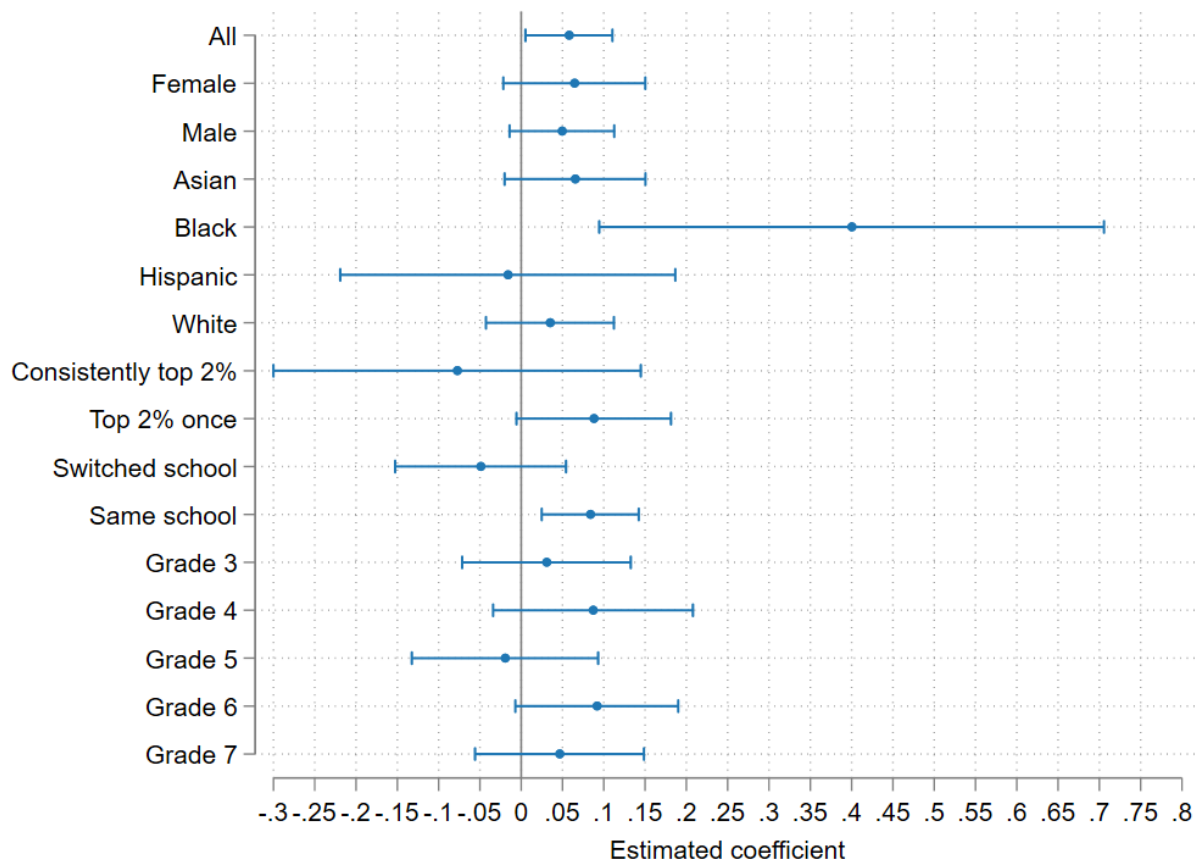


**Figure B4. DDD Estimates by Student Subgroup**



*Note:* Dots are the DDD estimates of the nudge effect in student standard deviation units, and whiskers represent the 95% confidence interval. The effects were estimated with student covariates as well as grade (except for grade-specific groups), year, and school fixed effects. Standard errors were clustered at the school level.

**Table B1. Comparison of Student Characteristics Between Math Top 2% Students and Other Students**

| Student characteristic            | Top 2%         | Other students  |
|-----------------------------------|----------------|-----------------|
| Next year's math score (SD units) | 1.94<br>(0.58) | -0.01<br>(0.97) |
| Prior math score (SD units)       | 1.89<br>(0.59) | -0.03<br>(0.97) |
| Disadvantaged                     | 0.10<br>(0.31) | 0.51<br>(0.50)  |
| English learner                   | 0.01<br>(0.09) | 0.10<br>(0.30)  |
| Low income                        | 0.07<br>(0.26) | 0.39<br>(0.49)  |
| Special education                 | 0.03<br>(0.17) | 0.20<br>(0.40)  |
| Male                              | 0.56<br>(0.50) | 0.51<br>(0.50)  |
| Asian                             | 0.35<br>(0.48) | 0.07<br>(0.25)  |
| Black                             | 0.02<br>(0.12) | 0.09<br>(0.29)  |
| Hispanic                          | 0.04<br>(0.19) | 0.22<br>(0.42)  |
| White                             | 0.52<br>(0.50) | 0.57<br>(0.49)  |
| Observations                      | 26,450         | 1,643,462       |

*Note:* Sample restricted to Grades 3–7, standard MCAS math takers in spring 2017, 2018, 2021, 2022, and 2023 with a subsequent-year MCAS record. Lagged scores further restricted to Grades 4–7 test takers in spring 2018, 2022, and 2023 (with 12,013 and 749,330 observations for the top 2% group and other students). “Disadvantaged” equals 1 if student is an English learner, economically disadvantaged, or in a special education program.

**Table B2. Difference-in-Differences Estimates: The MSAA Effect on Next Year's Mathematics Scores, Top-2% Sample**

|                      | (1)                | (2)                | (3)                | (4)                | (5)                |
|----------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                      | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023) | 0.16***<br>(0.01)  | 0.14***<br>(0.01)  | 0.06***<br>(0.01)  | 0.15***<br>(0.01)  | 0.09***<br>(0.01)  |
| Disadvantaged (T)    | -0.11***<br>(0.02) | -0.11**<br>(0.06)  | -0.18**<br>(0.07)  | -0.06<br>(0.06)    | -0.17**<br>(0.08)  |
| Post × T (DiD)       | 0.07***<br>(0.03)  | 0.07***<br>(0.03)  | 0.09**<br>(0.04)   | 0.05<br>(0.03)     | 0.08**<br>(0.04)   |
| Lagged score         |                    |                    | 0.29***<br>(0.01)  |                    |                    |
| English learner      |                    | -0.17***<br>(0.06) | -0.02<br>(0.08)    | -0.17***<br>(0.06) | -0.11<br>(0.08)    |
| Low income           |                    | -0.02<br>(0.06)    | 0.08<br>(0.07)     | -0.05<br>(0.06)    | 0.03<br>(0.07)     |
| Special education    |                    | 0.05<br>(0.05)     | 0.11*<br>(0.06)    | 0.03<br>(0.05)     | 0.09<br>(0.06)     |
| Male                 |                    | 0.02**<br>(0.01)   | -0.01<br>(0.01)    | 0.01<br>(0.01)     | 0.02**<br>(0.01)   |
| Asian                |                    | 0.12***<br>(0.02)  | 0.09***<br>(0.02)  | 0.11***<br>(0.02)  | 0.12***<br>(0.02)  |
| Black                |                    | -0.10**<br>(0.04)  | -0.08<br>(0.06)    | -0.10**<br>(0.05)  | -0.09*<br>(0.05)   |
| Hispanic             |                    | -0.14***<br>(0.03) | -0.09**<br>(0.04)  | -0.13***<br>(0.03) | -0.17***<br>(0.04) |
| White                |                    | -0.10***<br>(0.02) | -0.07***<br>(0.02) | -0.10***<br>(0.02) | -0.10***<br>(0.02) |
| N                    | 25,374             | 25,374             | 11,581             | 20,225             | 15,215             |
| R <sup>2</sup>       | 0.18               | 0.20               | 0.30               | 0.21               | 0.24               |

*Note:* All models include grade, year, and school fixed effects. Sample restricted to students who scored in the top 2% in math in grades 3–7. Standard errors clustered at the school level in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

**Table B3. Triple-Difference (DDD) Estimates: The MSAA Effect on Next Year's Mathematics, Top-2% Sample**

|                                  | (1)                | (2)                | (3)                | (4)                | (5)                |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                                  | No cov.            | With cov.          | Lagged score       | Pre=2017, 18       | Pre=2021           |
| Post (2022 and 2023)             | 0.09***<br>(0.00)  | 0.09***<br>(0.00)  | 0.03***<br>(0.01)  | 0.09***<br>(0.00)  | -0.10***<br>(0.00) |
| Disadvantaged (T)                | -0.21***<br>(0.00) | -0.03***<br>(0.01) | -0.02***<br>(0.01) | -0.03***<br>(0.01) | -0.06***<br>(0.01) |
| Top 2% (Z)                       | 1.09***<br>(0.01)  | 1.04***<br>(0.01)  | 0.63***<br>(0.01)  | 1.10***<br>(0.01)  | 0.91***<br>(0.01)  |
| Post $\times$ T                  | 0.02***<br>(0.00)  | 0.03***<br>(0.00)  | 0.03***<br>(0.01)  | 0.02***<br>(0.00)  | 0.05***<br>(0.01)  |
| Post $\times$ Z                  | 0.06***<br>(0.01)  | 0.05***<br>(0.01)  | -0.02*<br>(0.01)   | -0.01<br>(0.01)    | 0.17***<br>(0.01)  |
| T $\times$ Z                     | 0.07***<br>(0.02)  | 0.07***<br>(0.02)  | 0.02<br>(0.03)     | 0.08***<br>(0.02)  | 0.06*<br>(0.03)    |
| Post $\times$ T $\times$ Z (DDD) | 0.06**<br>(0.03)   | 0.06**<br>(0.03)   | 0.09**<br>(0.04)   | 0.05<br>(0.03)     | 0.06<br>(0.04)     |
| N                                | 612,501            | 612,501            | 287,291            | 516,842            | 323,343            |
| R <sup>2</sup>                   | 0.25               | 0.26               | 0.47               | 0.26               | 0.28               |

*Note:* Z = 1 for students in the top 2%; Z = 0 for students meeting expectations who are therefore not eligible to receive MSAA. All models control for the same student characteristics and fixed effects as specified in model (2) in Tables B2. Standard errors clustered at the school level in parentheses. \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.