Getting Down to FACTS

Adoption Windows and Reform: California's Math Pathways in the Common Core Era

Elizabeth Huffaker
University of Florida

May 2026

Stanford | SCALE Initiative
Accelerator for Learning

Getting Down to FACTS

Adoption Windows and Reform: California's Math Pathways in the Common Core Era

Elizabeth Huffaker

University of Florida

Executive Summary

When California districts faced the Common Core adoption window a decade ago, more than half reorganized their high school math sequence from Algebra–Geometry–Algebra II to an integrated model (IM), and even more curtailed access to algebra (or its integrated equivalent) in 8th grade, rolling back a generation of "Algebra for All" policy ("rollback"). This study examines the consequences of those choices, providing quasi-experimental evidence relevant to decisions districts continue to grapple with as they respond to the 2023 California Mathematics Framework.

Districts that adopted integrated math saw small improvements in 11th grade math achievement of roughly 0.037 standard deviations on the state assessment. This finding is sensitive to pre-trend violations and may not be the sole result of the integrated sequence itself: when math-specific growth is separated from concurrent improvements in English Language Arts, the IM-attributable effect becomes small and imprecise. A more defensible interpretation, consistent with the broader literature, is that enthusiastic CCSS adopters saw small, generalized learning gains, of which integrated math adoption may have been one component.

IM-adopting districts also saw slower student progression through the core high school sequence: the cohort-tracked share reaching precalculus and above declined by approximately two percentage points in standalone models. When IM and rollback are estimated jointly (i.e., through TWFE interaction models, CSDID subgroup decompositions, and controls for the other reform) rollback emerges as the primary driver of course-taking declines, while the IM-specific contribution is smaller and less certain. Districts that adopted IM alongside rollback actually saw smaller declines in advanced course-taking than districts that rolled back alone, suggesting IM may be mildly protective of student pacing. Notably, IM-associated pacing changes did not extend to the most advanced courses: calculus,

statistics, and AP/IB enrollment were unaffected. Slower pacing also does not explain the test score effect, whatever drove the small improvements likely operated through integrated (or otherwise CCSS-aligned) instruction itself.

The rollback of 8th-grade algebra access also reduced enrollment in precalculus and beyond. Defining a singular "rollback effect" is conceptually and empirically vexed, so impacts are documented across a range of binary cutpoints, a continuous-treatment estimator, and alternative control groups; estimates converge on a roughly −2 to −3 percentage point reduction in advanced course-taking among (i.e., 5 to 8 percent below baseline) districts that sharply limited 8th grade algebra. This effect is identified from a roughly 33 percentage point differential reduction in 8th-grade Math 1 enrollment between treated districts (64% to 10% between 2012 and 2018) and partially-treated comparison districts (67% to 46%). In other words, a roughly 1 percentage point decline in advanced course-taking corresponds to a 10 percentage point decline in 8th-grade algebra access. These effects are concentrated among the small-to-mid-size districts that make up most of the state California districts, suggesting larger districts may be able to buffer the downstream effects of deceleration through alternative pathways. Unlike IM, rollback reduced enrollment in the most advanced courses: calculus enrollment declined by roughly 1.2 percentage points. While modest in absolute terms, this represents roughly a 15 percent decline against baseline calculus enrollment in rollback districts.

In the statewide cohort-tracked data, the share of 11th- and 12th-grade students in advanced math courses fell about five percentage points between the 2015 and 2017 cohorts. About 60 percent of this decline is attributable to the IM and rollback adoption decisions; the remainder likely reflects a combination of broader secular trends and reform effects operating through high-school-only districts that are not directly observed in the analytic sample.

However, slowing the pace at which most students could move past middle-school coursework did not improve 11th grade test scores, and may have even reduced them in districts without the buffer of IM-associated learning. As expected, rollback reduced the share of 9th graders enrolled in Algebra I as repeaters. But it reduced the share of students still completing foundational coursework (Algebra I, Math 1) in 11th grade by less than IM and not at all in 12th grade. Furthermore, neither policy achieved the equity gains in the manner many reform districts intended. To the extent that these reforms narrowed racial gaps in advanced course access, it was through leveling down rather than lifting up:

advanced course take-up fell for students from groups typically more represented in advanced courses, while students from underrepresented minority groups (URM) saw no offsetting gains.

These findings carry direct implications for California districts navigating the current California Mathematics Framework adoption window. The full report develops five implications for the current California Mathematics Framework adoption window. Three are highlighted here. First, curricular adoption windows can be powerful catalysts for structural change. If motivated, district administrators are well-positioned during these moments to implement structural changes (i.e., course sequencing, placement thresholds) that directly shape which math students can access and when. This contrasts with the frequently diluted mechanisms of standards-based reform that depend on teacher-initiated changes to classroom practice. The scale and speed of the post-Common Core transformation underscore the leverage these windows carry.

Second, while the rollback of middle-school algebra was motivated by equity and achievement goals, the evidence suggests broad middle-school deceleration does not meaningfully support either goal in the spirit intended. Supporters hoped that deferring early algebra access would narrow attainment disparities by reducing differential acceleration and Algebra I repetition rates, and that prolonging student engagement with pre-algebra material would build a deeper mathematical foundation that paid off later. As this report documents, neither materialized. Deceleration did not improve learning, and where the racial gap in advanced course-taking did narrow, it did so by lowering the ceiling rather than raising the floor. These findings point toward more nuanced approaches (e.g., automatic enrollment) that avoid all-or-nothing placement and reduce reliance on the discretion of traditional gatekeepers.

Finally, the context behind the magnitude of these findings (e.g., the 1.2-point calculus effect amounting to a 14 percent decline from baseline) is a reminder that policy debates over acceleration focus intense attention on the relatively small, though important, group of students who ever reach calculus, AP math, and other advanced courses. If California's goal is to broaden the population of students prepared for advanced coursework while also strengthening the mathematical development of all students — regardless of whether they ever reach these classes — attention must be focused

primarily on the greater challenge of improving instructional quality in elementary and middle school grades.

Acknowledgements – I would like to thank the members of my dissertation committee — Thomas Dee, Susanna Loeb, Eric Bettinger, Sarah Cohodes, and Jennifer Osuna — for their guidance on this research, with particular thanks to Tom and Susanna for their ongoing support of this project. Thank you also to Morgan Polikoff for his comments on an early conference paper and Chris Doss for his detailed and insightful feedback on a recent version of this report. I am grateful to Ashley Campbell, Alix Gallagher, Jana Luft, and Lisa Towne for conducting and sharing the superintendent interview data, the Getting Down to Facts III team including Nancy Waymack and Julia Perlmutter as well as Ev Fu Gilbert and Briana Jocelyn for their contributions to the principal qualitative data collection.

Introduction

The structure of secondary math course sequences shapes who has access to advanced coursework in high school and, by extension, who is most prepared for quantitative fields in postsecondary education and the labor market. Because advanced math course taking is among the strongest predictors of college enrollment and completion (e.g., Adelman, 2006; Goodman, 2019; Long et al., 2012), decisions about how districts organize math pathways—which courses are offered, in what order, and to whom—carry consequences that extend well beyond the math classroom. Simultaneously, advanced course taking patterns have long been vexed by stubborn ethnoracial and socio-economic disparities (Conger et al., 2009).

California's attempts to balance these achievement and equity considerations in structuring secondary math opportunities have driven national conversations throughout this century. In the 2000s, they led the nation in encouraging broad expansion of middle school access to high school coursework under the "Algebra for All" movement (Domina et al., 2015). In the 2020s, a revised advisory framework sparked heated discussion over the merits of broadly *deferring* Algebra I until 9th grade, a reversal some have dubbed "Algebra for None" (e.g., Luna, 2023). The missing link to understanding the dramatic pendulum swing between these positions is a comprehensive account of the policy context and consequences of the intervening decade, which coincided with the adoption of the Common Core State Standards (CCSS).

In this report, I show how the CCSS rollout opened a policy window that reshaped California's secondary math pathways and provide quasi-experimental estimates of resulting impacts on students' high school trajectories. Specifically, while California is a state with historically centralized control over math education, its implementation of the CCSS devolved key pathway decisions to districts. Within a few years, more than half of California's districts had reorganized their high school math sequence (i.e., adopting "integrated" math (IM) that weaves algebra, geometry, and statistics into each year rather than teaching them as separate courses), rolled back middle school access to algebra ("rollback"), or both. This produced a fragmented policy landscape where statewide uniformity had previously prevailed. In the ensuing decade, advanced secondary math enrollment has declined among California students (Figure 3; Dykeman et al., 2026). This paper leverages within-state variation in pathway

reforms take-up and timing to present estimates of the extent to which observed changes in course taking and student achievement were plausibly driven by specific CCSS-era policy choices.

Study Overview & Research Questions

California offers unusual analytic leverage for studying these dynamics. The scale of CCSS adoption generated wide variation in both the timing and intensity of pathway reforms across over 300 districts, observable in administrative data from the California Department of Education (CDE) through the 2018-19 school year, after which the CDE has not made course enrollment data publicly available. Because districts adopted these reforms at different times, the staggered nature of treatment supports a difference-in-differences (DiD) design that can help distinguish reform-associated changes from common time trends. The relevance of this analysis extends beyond the historical record as California is currently in a new curriculum adoption window following the 2023 California Mathematics Framework. Districts are again making consequential decisions about pathway structure yet have little empirical evidence from the prior reform window to guide them.

To address this gap, I draw on district-level panel data covering 395 districts from 2012 to 2018, using the Callaway and Sant'Anna (2021) DiD (CSDID) estimator to accommodate staggered treatment timing. I complement the quantitative analyses with interviews with 35 principals and 94 superintendents across 69 California districts conducted in 2025.

This report addresses the following research questions:

1. How did the California secondary math course pathways landscape change following the adoption of the Common Core State Standards?
2. What were the effects of integrated math adoption on student math achievement and course taking?
3. What were the effects of the rollback of middle school access to high school math on student math achievement and course taking?
4. To what extent did these reforms shape ethnoracial disparities in math achievement and advanced course access?
5. How are district leaders navigating math pathway decisions during the current cycle of curriculum adoption?

I examine course taking along several dimensions. The primary focus is advanced math attainment: the share of students progressing beyond the core three-year sequence into precalculus, calculus, statistics, and AP/IB mathematics. I also examine whether reforms affect lower-level course repetition and overall math enrollment intensity in 11th and 12th grade.

Preview of Results

This report makes several contributions. First, it provides quasi-experimental evidence on two prominent math education policy debates with direct relevance to decisions California district leaders face today. Second, it introduces a novel perspective for conceptualizing and measuring CCSS effects by examining policy decisions catalyzed, though not required, by the standards. Finally, it demonstrates how standards adoption more generally can function as a powerful policy window for ancillary structural reforms.

I find that districts saw small (ES = 0.037 SD, $p<0.01$) – and pre-trend sensitive – improvements in 11th grade math achievement on the Smarter Balanced state assessment (SBAC) following the adoption of IM. However, when I attempt to separate IM-specific gains from concurrent improvements in district achievement shared across math and ELA, the IM-attributable effect becomes small (0.005 – 0.014 SD) and imprecise. This is not dispositive of the benefits of IM adoption (i.e., districts could have simultaneously adopted highly effective ELA programming), but a more defensible interpretation, consistent with the literature (Bleiberg, 2021; Polikoff, 2017), is that enthusiastic CCSS adopters saw small, generalized learning gains. Notably, controlling for 11th grade course composition does not diminish the SBAC effect: IM districts did not improve learning by changing the pace at which students completed the core high school sequence (Table A4).

The second policy considered, the decline in middle school Algebra I (or Integrated Math I, collectively "algebra") access ("rollback") produced no detectable changes to student learning as measured by the SBAC in my main specification, but the direction of the point estimate is consistently negative across specifications and significant when simultaneously estimated with IM effects.

In terms of course-taking, using my primary standalone specification, I estimate that IM adoption was associated with reductions in the combined cohort-tracked share of students enrolled in

advanced math courses by approximately two percentage points, driven by reduced precalculus[1] take-up. This measure, which pools 11th and 12th grade enrollment for the same cohort, reflects the net effect on a cohort's advanced course taking trajectory (i.e., it absorbs mere shifts in course timing). The reduction is driven by a larger shift at 11th grade that partly attenuates by 12th grade.

Reduced access to high school level coursework was also associated with reduced average advanced course taking. The cohort-tracked measure shows a slightly larger combined decline of nearly three percentage points. Because rollback is a continuous, time-varying treatment that leading staggered DiD estimators handle imperfectly, I document this effect across multiple binary cutpoints, a continuous-treatment estimator, and alternative control groups, with estimates converging on the −2 to −3 point range (i.e., a 5 to 8 percent drop). Unlike the IM effects, the rollback-driven course-taking decline shows limited attenuation between 11th and 12th grade and reduces calculus enrollment by slightly over one percentage point across specifications. When both reforms are estimated simultaneously, rollback emerges as the primary driver of course-taking declines and IM independently contributes a smaller but statistically significant effect.

Both reforms are associated with narrower racial gaps in advanced course access, but through leveling down rather than lifting up: Non-URM students lost more access to advanced courses than URM students relative to comparable non-adopting districts. The cohort-tracked gap narrowed by roughly 2.5 to 3.2 percentage points (i.e., a 25 to 35 percent reduction relative to the 9.1-point pre-reform gap). I find some evidence that IM supported lower achieving students. It reduced the share still enrolled in introductory coursework in 11th and 12th grade and shifted roughly 1.4 percentage points of students across the state's math proficiency threshold. Algebra rollback did lower first-time repetition rates in ninth grade, but reduced the share of students enrolled in lower-level courses towards the end of high school by less than IM (i.e., just over 1 pp). There is no evidence that sharply reducing 8th grade algebra enrollment produced average benefits for learning or advanced course-taking for groups the policy was often designed to serve.

The cohort-tracked share of students in advanced math courses peaked at roughly 37 percent for the 2015 cohort before falling to about 32 percent for the 2017 cohort — a decline of close to five percentage points in the unified sample. Weighting each treatment subgroup's estimated effect (Table

[1] Includes Integrated Mathematics IV course code.

2, Panel D) by its share of the sample, the IM and rollback adoption decisions together account for a majority of this decline. What remains may reflect reform effects in high-school-only districts whose feeder-district policies are unobserved here, or a secular deprioritization of acceleration.[2]

Policy Context and Adoption Landscape

Window 1: CCSS Adoption

California adopted the Common Core State Standards in August 2010, but full implementation unfolded over several years. The State Board of Education approved a new mathematics framework in 2013, which provided guidance on curriculum, instruction, and the organization of high school course sequences. The state completed its transition from pre-CCSS aligned California State Tests (CSTs) with the first full administration of the Smarter Balanced assessments (SBAC) occurring in AY 2014-2015.

This extended timeline is reflective of the circuitous path by which the national Common Core State Standards movement aimed to influence classroom practice and ultimately student learning. The theory of action for standards-based reform rests on the logic that common rigorous standards can align the broader instructional system of assessments, curriculum materials, and teacher preparation around shared expectations for what students should be able to know and do, tightening the link between what the state expected and what teachers taught (Polikoff, 2017). However, they aim to impose this coherence on a fragmented school system without fundamentally challenging the localism that characterizes American school governance. The translation of the standards into practice therefore depends on a cascade of decisions at the state, district, school, and teacher level (Loveless, 2021).

Yet the CCSS theory of action was fundamentally about shifting the instructional core – teaching and learning inside the classroom (Elmore, 2004). That ambition is vulnerable to exactly this kind of implementation chain. The historical record of standards-based reform is one of dilution. As Cohen and Ball (1990) observe – and Loveless (2021) applies to the CCSS case – policies whose theory of action runs through shifting teaching are challenged because “teachers are at once the targets and the agents of change.” The underwhelming evidence on aggregate CCSS effects is consistent with this analysis

[2] Because these estimates are derived from the Unified districts for which both treatments are observed, the projection to the full statewide trend is approximate.

(Polikoff, 2017). Empirical retrospectives show modest gains from standards adoption itself (Bleiberg, 2021; Gao & Lafortune, 2019; Xu & Cepa, 2018) and small returns, without evident generalizability, from aligned professional development (Allensworth et al., 2021). Despite the prominent scope and cost of the CCSS, the instructional levers it was designed to activate have proven characteristically resistant to external direction (Elmore, 2010; Tyack & Cuban, 1995).

The pathway reforms examined in this report operate through different institutional mechanisms. Despite comprising one of the more visible legacies of the CCSS era, especially in California, these policies were merely legitimated, not mandated, by the standards themselves. Specifically, the CCSS-M appendix presented both integrated and traditional sequences for core high school coursework (CCSS Initiative, 2010). The standards also described accelerated pathways to eighth-grade Algebra I but cautioned against premature acceleration from the default three-year middle school sequence. This equivocation is consistent with the CCSS architects' broader effort to maintain fragile compromise across diverse stakeholders (Loveless, 2021). According to CDE CALPADS data, divergence from the pre-CCSS status quo proved popular: among Unified districts (K-8 districts are not observed in this dataset), the share of eighth graders in high school math fell from about 65% in 2012 to 20% in 2018, while the share of ninth graders in an integrated sequence among all high school-serving districts increased from essentially none to almost half by 2018.

Three features may help explain the broad take-up of IM and rollback. First, unlike reforms that depend on changes to classroom practice, IM adoption and rollback operate through channels less susceptible to the dilution that characterizes instructional reform (Loveless, 2021; Pressman & Wildavsky, 1973). Decisions for how to structure students' educational experiences (i.e., what courses they took and when) are under district rather than teacher control. These policy choices influence teachers' instructional conditions, such as the scope and sequence of their classes, and the composition and preparedness of their students. But, their influence on student outcomes does not inherently depend on teacher practice – administrators are the agents of change. A related but weaker reform of this nature is textbook adoption, where state and district decisions can influence student learning (e.g., Koedel et al. 2017) by changing the instructional context teachers operate within. Even so, teachers have more meaningful autonomy over the usage of textbooks than they do over the course placement decisions of students and the macro-structure of the courses they teach.

Second, by locating these decisions at the district level, their immediate political salience was muted. California historically exercised considerable state-level control over math curriculum including statewide textbook adoption until 2012 and accountability mechanisms that encouraged eighth-grade algebra enrollment (McEachin et al., 2020). This heavy state-level hand plausibly created the staging ground for the 1990s "Math Wars" (Becker & Jacob, 1998). By contrast, the widespread rollback of early algebra in the 2010s drew sustained blowback only in districts with significant local media presence or highly engaged parent populations (e.g., SFUSD, Tintocalis, 2015; San Mateo Union HSD, Loveless, 2021). It is unclear whether the scale of either reform (Figures 2, B1) is well understood even now. Media coverage tends to frame the removal of middle-school algebra as the pathology of a few unusual districts when roughly one hundred unified districts reported essentially zero 8th-grade Algebra I or Mathematics I enrollments in 2018.[3]

Finally, the CCSS legitimized a menu of clearly defined policy options for which there was likely latent demand and eager policy entrepreneurs ready these specific reforms with those needs (Kingdon, 1984). As Binder (2002) showed in other curricular domains, adoption cycles can function as policy windows in which reforms that appear to address recognized institutional problems gain traction through bureaucratic channels, even when they extend beyond the original impetus for change. For example, when SFUSD Superintendent Carranza proposed removing Algebra I from 8th grade, he cited the CCSS and the need to not rush students through material but more so emphasized discussing the equity motivations of his plan, responding to highly disproportionate failure rates that characterized the “Algebra for All” era (Carranza, 2015).

In sum, while the CCSS granted states considerable latitude over implementation and positioned teachers as the ultimate agents of change, the pathway decisions examined here elevated a different set of actors. District administrators became key decision-makers within California’s implementation pyramid, and the structural choices they faced (i.e., integrated or traditional, accelerate or decelerate) were sufficiently narrow and well-defined to produce rapid, coherent reform at scale. The result was a

[3] This count is restricted to California unified and secondary districts that offered Algebra I or Mathematics I at scale (i.e., more than 20% of 8th graders enrolled) in 2012 and “essentially zero" is defined as 2% or less. Enrollments coded as "Other" (CDE's residual 8th-grade math category) are not counted as algebra; under an adversarial bound that reclassifies every "Other" enrollment as Algebra I or Mathematics I, the count falls to 70. However, this bound over-adjusts as known cases (e.g., SFUSD) that code replacement course as "Other" and would be mechanically excluded under this rule.

fertile policy arena for mathematics approaches that seem, at least superficially, distal to the central scope and aims of a delicately calibrated common standards movement.

Policy Decision 1: Integrated vs. Traditional Math

In most American high schools, math courses cover discrete content areas (i.e., Algebra I, Geometry, Algebra II) in a sequence commonly referred to as the "traditional" (or AGA) pathway. This model stands apart from both international programs of study and the organization of K-8 mathematics and has been critiqued by math education scholars for decades (Shaughnessy, 2011a). An alternative, “integrated sequence,” weaves content from algebra, geometry, and statistics across each year of high school. Proponents identify several theoretical mechanisms through which this approach may generate superior learning outcomes. Integration of topics across mathematical domains within the same school year can facilitate richer inquiry into real-world relationships, improving both algebraic and geometric problem-solving (Okano & Tsuchiya, 1999). An integrated sequence may also better reinforce key concepts through spaced repetition across the student's high school trajectory, a principle well supported by cognitive and educational psychology research (Kang, 2016; Noh & Webb, 2015). And teachers working within an integrated system may themselves develop stronger pedagogical content knowledge by deepening their understanding of the links across core concepts (Noh and Webb, 2015; Shulman, 1986).

Beyond these instructional mechanisms, the introduction of an integrated sequence could promote flexibility in the opportunity structure of students' math course taking. The prevalence of tracking in high schools is critiqued in part for the high degree of rigidity across tracks (Domina et al., 2019). A student intending to take calculus without doubling up or attending summer school must already be in Geometry by the time they enter 9th grade, and moving across tracks within an AGA system is further constrained by the course organization of math content. An integrated sequence could more flexibly prepare students for a wide range of advanced options after the three-course course sequence.

The state of existing empirical evidence on the efficacy of an integrated sequence compared with the traditional pathway is limited. Advocates often point to the superior performance of countries like Japan on international assessments (Galley, 2004), but the role of integrated math in deriving this

correlation is not supported by credible confirmatory analysis. The absence of domestic evidence reflects the rarity of the integrated approach in the United States prior to the CCSS (Will, 2014). In general, causal studies of high school math policies are challenged by data limitations and a highly fragmented policy landscape (Canaan et al., 2022; Han et al., 2023). At present, the most substantial integrated math quantitative study compared student outcomes within schools where both traditional and integrated courses were offered, finding that students in integrated pathways outperformed their peers on a variety of measures (Confrey and Maloney, 2015; Tarr et al., 2013). However, because students self-selected into their pathway, these estimates cannot be interpreted as strictly causal.

The CCSS created a policy window to disrupt the AGA paradigm when it outlined recommended pathways for both traditional and integrated sequences, lending formal legitimacy to the integrated model for the first time (Shaughnessy, 2011a; Will, 2014). Still, there are practical and political barriers to implementing integrated curricula in the United States. Skeptical administrators have pointed to the lack of curricular materials, considerable parent and teacher opposition, and the challenge of retraining teachers as reasons to retain the traditional sequence (Confrey & Maloney, 2015; Harlow, 2015). These concerns may explain why such reforms were sparse prior to the Common Core (Will, 2014) and why Georgia – the only state that had previously used an integrated sequence – returned to the AGA pathway in 2015 following CCSS-associated backlash (Loewus, 2015). Despite these risks, the CCSS ignited a wave of sequence reform across several states and many districts. This study provides an opportunity to measure the effects of this wide-scale curricular shift.

Policy Decision 2: Early Access to High School Math

Integrated math adoption occurred alongside a second distinct structural shift in California's math pathways: the rollback of early access to high school math. This shift had three main drivers. First, a body of quasi-experimental evidence suggests that the "Algebra for All" movement of the 1990s and 2000s had not delivered on its promise of broadening access to advanced math and may have actually harmed the mathematical advancement of lower-proficiency students (Clotfelter et al., 2015; Domina et al., 2015; McEachin et al., 2020). Beginning in the late 1990s, a series of California policy initiatives expanded eighth-grade algebra I enrollment from 16% in 1999 to a peak of over 60% by AY 2012-13 (Domina et al., 2015; Rosin et al., 2009). The case for universal early access rested largely on

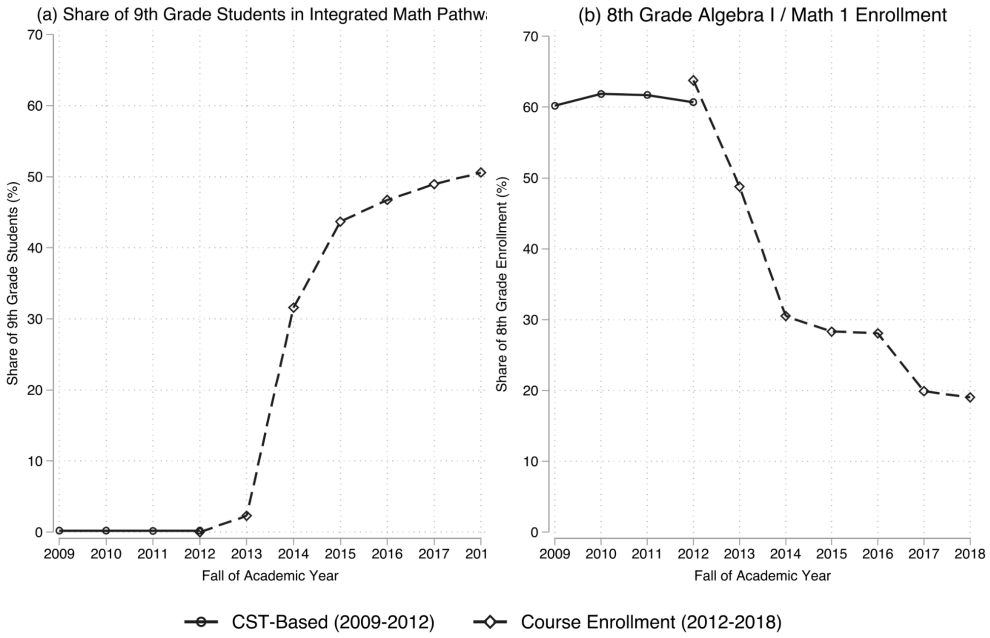
observational findings associated with eighth-grade algebra enrollment (Gamoran & Hannigan, 2000; Stein et al., 2011) and on documented disparities in which minoritized students were substantially less likely to be placed in algebra before 9th grade (Conger et al., 2009; Oakes et al., 1990). However, as "Algebra for All" policies scaled, regression discontinuity studies found that acceleration often carried negative consequences for students lacking prerequisite skills and knowledge, while producing some gains for high performers (Clotfelter et al., 2015; Dougherty et al., 2017; McEachin et al., 2020). In California, the broad expansion of 8th-grade algebra was associated with overall declines in 10th-grade math achievement (Domina et al., 2015).

Second, the CCSS introduced a more rigorous three-year middle school math sequence that did not presuppose acceleration into high school content before 9th grade. California's guidance specifically cautioned that decisions to accelerate students into higher mathematics before 9th grade should not be rushed (Fensterwald, 2014). The promise of "depth over breadth" embedded in the CCSS middle school math pathway was appealing in this context – perhaps slowing the pace at which students proceed through foundational content would reap learning rewards in the long run (Boaler et al., 2018). Third, some districts framed the rollback in explicitly equity-oriented terms, citing racially stratified repetition rates and advanced course-taking patterns as evidence that early acceleration had entrenched, rather than reduced, inequality (Carranza, 2015; Huffaker et al., 2025). "Detracking" strategy shifted from an aspirationally universal model of acceleration to one of deceleration.

The CCSS Reform Landscape

Before turning to the contemporary curriculum adoption window, I answer RQ1 by documenting the scope and character of pathway changes that emerged during the Common Core era.

Figure 1. Policy Context: IM Adoption and 8th Grade Algebra Rollback Trends



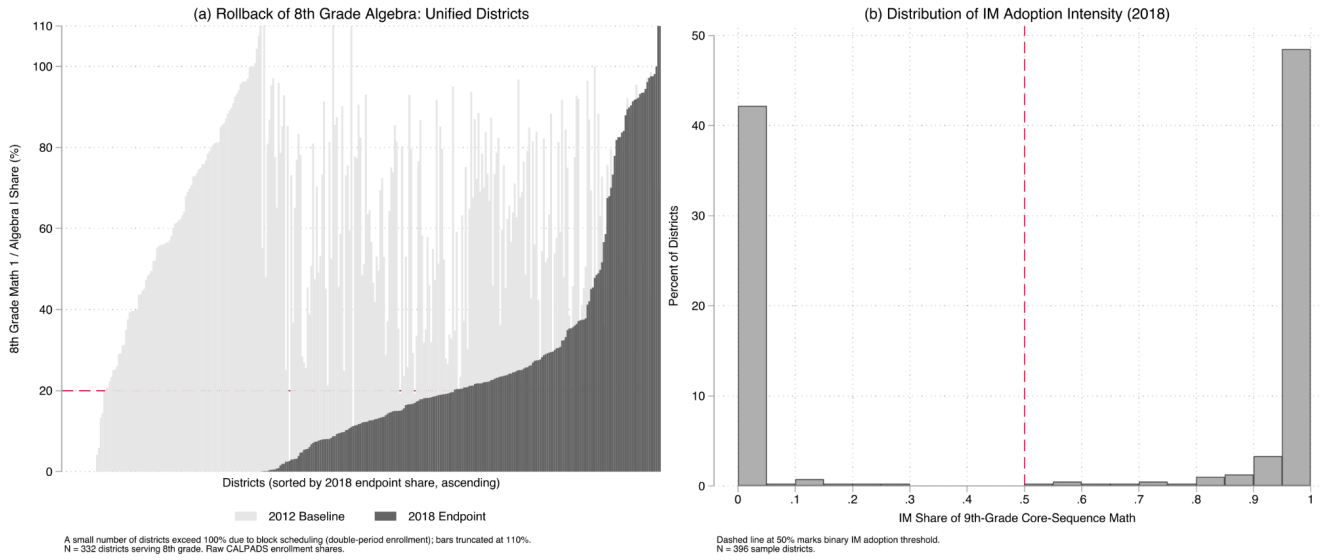
Notes: Panel A plots the enrollment-weighted share of 9th graders in Integrated Math courses, drawn from the 395-district sample that reports state test data across the panel (330 districts serving 8th grade plus 62 high school districts). Panel B plots the enrollment-weighted share of 8th graders in Math 1 / Algebra I at Unified and high school districts serving 8th grade. Course enrollment data (dashed line with diamonds) come from CALPADS; the CST assessment files (solid line with circles) provide the pre-CALPADS measure.

Figure 1 Panel A presents state-level trends in both reforms across the 395 traditional high school serving districts in my sample using course enrollment data (2012-2018) from the CDE's CALPADS directory and supplemented by California State Test assessment files through 2012 (Figure A1). It shows the share of ninth graders enrolled in an integrated math pathway. Prior to AY 2013-14, the share was effectively zero according to both data sources available (see Figure A1). Adoption then accelerated rapidly. By 2014, 37% of districts (32% of students) had adopted an integrated pathway. By 2015, 49% (44% of students), and by 2018, 56% (51% of students). The largest single cohort of adopters, 122 districts, transitioned in 2014, the first full year of CCSS implementation. Panel B features the subset who also serve 8th grade (N=330) and shows the concurrent decline in eighth-grade algebra enrollment. The share of 8th graders enrolled in algebra 1 or its integrated equivalent (Mathematics 1)

peaked at roughly 65% in AY 2012-13.[4] By 2018, the share had fallen below 20%, a decline of over 40 percentage points or two thirds, in six years.

These state-level aggregates obscure the extent of reshaping at the district level. Although media attention on post-CCSS delays to Algebra I has focused on a handful of high-profile districts, the rollback of middle school access to high school math was in fact widespread. Figure 2, Panel A presents a skyline plot of 8th-grade algebra enrollment for each of the 330 unified districts in the analytic sample, with light bars showing 2012 baseline shares and dark bars showing 2018 endpoint shares. Of the 213 districts where a majority of 8th graders were enrolled in Algebra I or its equivalent at baseline, 128 (60 percent) had reduced this share to below 20 percent by 2018, with a median decline of 73 percentage points and a median endpoint below 1 percent. Twenty-eight percent of unified districts (N = 91) dropped from a majority of 8th graders taking algebra in 2012 to fewer than one in ten by 2018. Supplemental Figure B1 shows deceleration occurred in all regions.

Figure 2. Policy Variation: 8th Grade Algebra Rollback and Integrated Math Adoption



Notes: Panel (a): Each bar represents one unified or 8th grade-serving high school district, sorted by 2018 endpoint share (ascending). Light bars show 2012 baseline; dark bars show 2018 endpoint; dashed line marks the 20% rollback-eligibility threshold. N = 330. Panel (b): Distribution of 9th-grade integrated math share across 395 sample districts in 2018. Dashed line at 50% marks the binary IM adoption threshold.

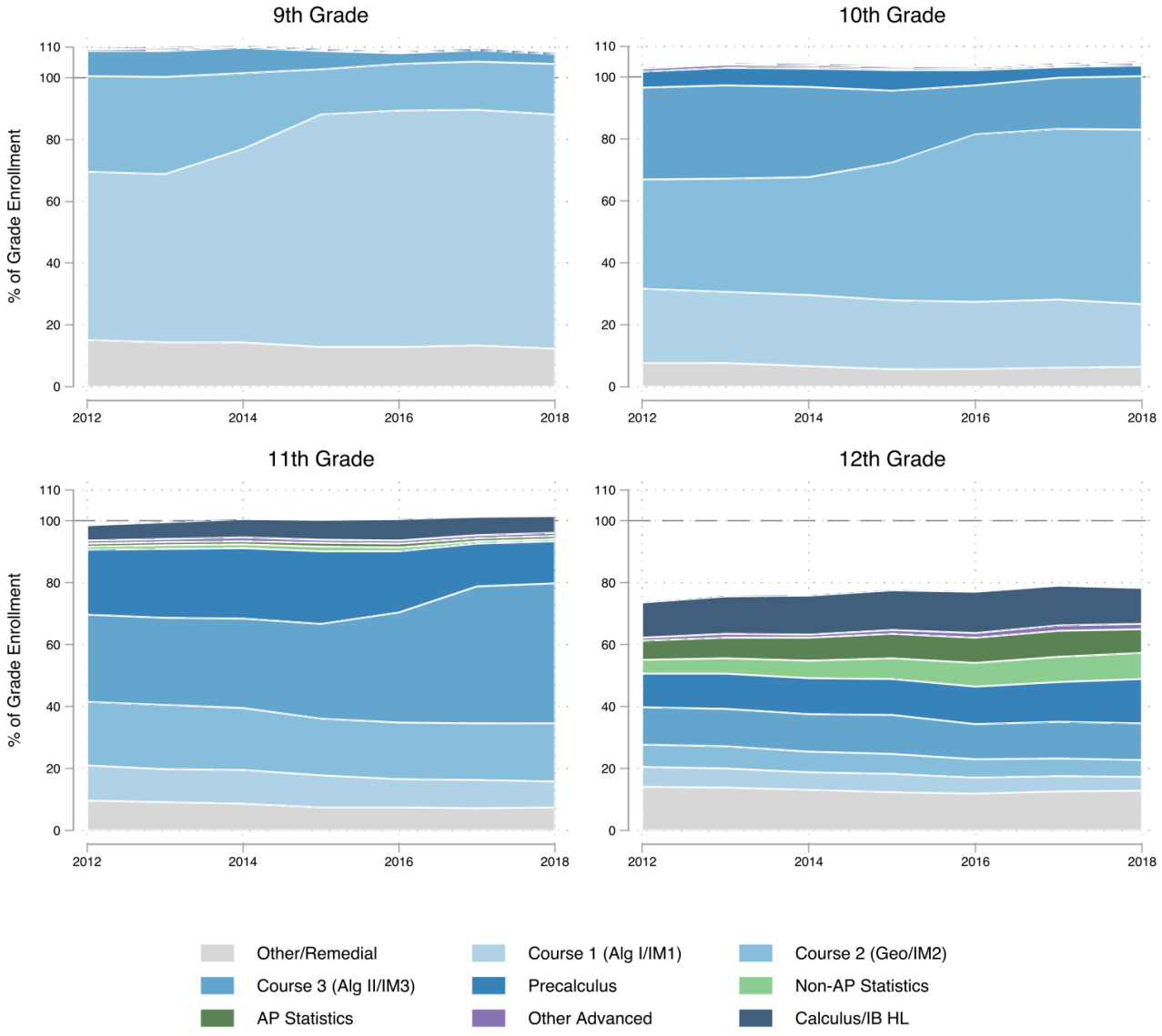
[4] CALPADS course data measure fall enrollment where the California State Test (CST) is taken in spring, which may explain the level difference.

Figure 2, Panel B shows the corresponding distribution of integrated math adoption intensity across all 395 sample districts in 2018 (see Data section for measure construction). The distribution is sharply bimodal: 161 districts (41 percent) had no 9th graders enrolled in integrated math, while 192 (49 percent) had at or near 100 percent of 9th graders in the integrated pathway. Only about 6 percent of districts occupied an intermediate position, consistent with integrated math adoption being a curricular decision made at the district level. Importantly, these two reforms were not independent. Figure B1 maps the geographic distribution of both policies and reveals substantial overlap: 75 percent of districts that adopted integrated math also rolled back 8th-grade algebra to below 20 percent, consistent with both decisions being prompted by a shared embrace of practices viewed to be aligned with Common Core best practice during the adoption window.

These policy shifts (i.e., IM adoption at 9th grade and rollback of 8th-grade Algebra I) overlapped with a visible reshaping of high school course-taking patterns that can be traced grade by grade. Figure 3 shows the evolution of course composition for grades 9 through 12 over the study period. The 45-percentage-point drop in 8th-grade Algebra I enrollment documented in Figure 1 propagates upward through the sequence on a predictable one-year lag. At 9th grade, Course 1 enrollment (Algebra I/IM I) rose from roughly 53 to 72 percent as displaced 8th-grade cohorts arrived to take algebra for the first time, while 9th-grade Geometry (or IM 2) fell by about 21 percentage points (i.e., 30 to 9 percent). Figure B2 presents the implied repeater rate of 8th-grade Algebra I students and confirms that rollback did reduce course repetition, consistent with rationale for the reform.

At 10th grade, traditional Algebra II enrollment fell by roughly 18 percentage points (from 29 to 11 percent), while combined Geometry and IM II enrollment rose by about 19 points as students moved into the on-time (rather than accelerated) Geometry position. At 11th grade, Course 3 (Algebra II / IM III) grew from 27 to 43 percent while precalculus declined by roughly 7 points (from 22 to 15 percent). At 12th grade, the deceleration arrives as a 9-point rise in precalculus enrollment (21 to 30 percent), with calculus and advanced electives largely unchanged. Because a fourth year of math is not required for high school graduation or to in-state university requirements, many students do not take 12th-grade math at all, and the mechanical compression observed in prior grades broadly dissipates. These raw trends motivate the analyses that follow: to what extent were course-taking shifts toward the end of the high school pipeline driven by IM adoption and/or the rollback of middle school algebra?

Figure 3. Math Course Composition by Grade Level as a Share of Enrollment



Notes: Course enrollment data from CALPADS (2012–2018). Enrollment shares exceed 100% because some students are enrolled in multiple math courses in a single grade level.

Policy Window 2: The New California Math Framework (CMF)

In 2021, just over a decade after California affirmed its adoption of the CCSS Math Standards, and eight years after its last framework revision, the California Department of Education released a draft version of a new California Mathematics Framework (CMF) structured around big ideas. It

contained recommendations that proved contentious, resulting in a protracted public review process. Among the draft's most debated features were recommendations to delay Algebra I until ninth grade for all (CDE, 2021). This contrasts with the 2013 CMF, which did not explicitly discourage entry to high school mathematics until high school (Loveless, 2023). The draft drew sharp criticism from a coalition of university mathematicians and parent advocacy groups who argued that it would restrict access to advanced math and preparation for quantitative college majors (Conrad, 2023; Fensterwald, 2021). Supporters, including math education researchers, countered that the framework reflected current evidence on the harms of premature acceleration and the benefits of deeper conceptual understanding (Boaler, 2016; Schoenfeld, 2022). After multiple rounds of public comment and substantial revision the State Board of Education adopted the final framework in July 2023. The final 2023 framework dropped the earlier draft's proposal for a common course sequence through tenth grade and explicitly acknowledged that "some students will be ready to accelerate into Algebra I or Mathematics I in eighth grade" (CDE, 2024).

In addition to drawing attention to the middle school acceleration debate, California's framework adoption may also impact on district decision-making about math pathways because it triggers a statewide instructional materials review and adoption window. Districts that adopt state-recommended materials receive dedicated funding, creating a financial incentive to revisit their math course sequences in alignment with the new framework's guidance (Gallagher et al., 2026). This materials adoption cycle, the first since CCSS implementation, represents a second policy window in which districts may reconsider decisions made a decade earlier regarding integrated vs. traditional sequences and the timing of algebra access.

The framework itself is advisory rather than mandatory, and districts retain full authority over course sequencing and placement decisions (CDE, 2015). While too recent to draw conclusions about its effects, the CMF could ignite a qualitative shift in the politics of pathway reform. During the CCSS transition, decisions about course sequencing and acceleration were made at the district level by superintendents, curriculum committees, and school boards acting within their traditional sphere of authority. Integrated math adoption drew some skepticism (e.g., Loewus, 2015) as did deceleration of Algebra I in some locations (e.g. SFUSD; Tintocalis, 2015) after implementation. However, most of these consequential CCSS-era pathway decisions drew little contemporaneous attention and were dwarfed

on the national stage by debate over the standards themselves and the role the federal government played in promoting them (Loveless, 2021). Just as the nationalization of previously local political issues can elevate conflict and polarization (Hopkins, 2018), the CMF effectively "state-ified" these decisions. By centering them in a statewide policy document, the salience and contestedness of choices that districts had already been making, often without controversy for a decade, was elevated.

The combination of updated state guidance, new instructional materials, and renewed public attention to math pathway policies creates conditions under which the structural decisions documented in this study may once again be in flux. While public administrative data is limited for the new adoption window,[5] this study captures present pathway policies using contemporaneous interview and survey data from district leaders. I also draw on open-ended interview responses from these administrators to document which factors (e.g., the new framework guidance, ongoing equity concerns, and stakeholder pressures) are shaping current placement and pathway deliberations.

Data & Sample

I primarily use publicly available datasets sourced from the California Department of Education (CDE) for the analyses reported in this paper. Below I describe these data, as well as the construction of key variables and the analytic samples.

Data Sources

Directory Information

I use CDE directory data to identify grade-span, charter status, alternative school status, and active status of schools and districts. These traits are used to refine my sample to the subset that serves high school students, as I describe further in subsequent sections. I supplement these with two district-level classifications from the NCES Common Core of Data: LEA type (used to identify traditional districts operating high schools) and locale (city, suburban, town, or rural).

[5] The CDE has not released publicly available aggregate course data more recent than AY 2018-2019.

Course-Taking

Course enrollment data for California public high schools are available from the CDE for AY 2012-13 through AY 2018-19. I use these data to construct treatment indicators (i.e., does the district offer an integrated or traditional pathway? Are 8th graders enrolled in high school level coursework?) as well as measures of student outcomes (e.g., the share of 12th graders enrolled in an AP class). Course enrollment data are available as school-grade-class-gender observations and are reported as both totals and by racial/ethnic category. I use district-cohort sums for my main analysis. This two-year total approximates cumulative upper division course taking for a given graduating class, though it is defined for fewer academic years due to the required lag. The CALPADS course enrollment files provide eighth grade enrollment in Algebra I or Integrated Math I. I use these data to construct my primary measure of early access to high school coursework and therefore to define Math 1 (i.e., Algebra I or Integrated Math I) rollback treatment status. I use raw enrollment shares despite the presence of some double counting and confirm through sensitivity analyses that imposing an adjustment for apparent double-period scheduling produces identical results (Table A5). Because eighth grade enrollment is only observed within the same administrative unit, this measure is available only among the sample (i.e., high school serving districts) for districts that operate both middle and high school grades (i.e., Unified K-12 and Union High 7–12 districts), covering 330 districts and three quarters of all California 8th graders in non-charter public schools.

I make a small number of manual adjustments following checks for quality and completeness. For eight district-year observations in my analytic sample, course data are reported for six of seven possible years (see technical appendix). I verify that the district is otherwise present in the sample (i.e., it is active in that year) and either impute a midpoint from the year before and after or use the neighboring observation for terminal data years. These imputations maintain the balanced panel required by the Callaway and Sant'Anna (2021) estimator; in no case do they fall between a change in treatment status, and dropping these districts yields substantively identical results. I also manually correct four district-year observations in which integrated and traditional course counts appear unintentionally inverted. Finally, I use Census Day enrollment totals to conduct quality checks on course totals (i.e., do total 9th-grade math enrollments roughly match grade-level enrollment). Because I am drawing from aggregate data, students enrolled in multiple math courses or a double period of the

same course are counted twice in the numerator. As such, many districts have math enrollment moderately exceeding 100%.[6]

I classify the approximately 60 unique CDE course codes according to Table B1. In addition to creating comparable course-level categories (e., Precalculus and Integrated Math IV are equivalent) into a six-category hierarchy reflecting progressive levels in the high school math pipeline: Course 1 encompasses first-year courses in either sequence (Algebra I, or Integrated Math I). Course 2 includes second-year courses (Geometry or Integrated Math II). Course 3 captures third-year courses (Algebra II or Integrated Math III). Course 4 groups Precalculus (or IM IV) and statistics (including AP Statistics) – courses typically taken at the same pipeline stage after the third-year course prerequisite. Course 5 includes calculus (AP Calculus AB and BC, non-AP calculus), IB Mathematics HL, and other advanced courses such as Linear Algebra. A Residual/Other remedial category captures prealgebra remedial courses and Financial Algebra, which is treated inconsistently across districts. In most cases, defining these labels is straightforward. For the small number of ambiguous codes (e.g., math analysis), I consulted the University of California A-G course lists and district websites to determine appropriate placement. Four districts (approximately 1% of the analytic sample) use an alternative Algebra I to Algebra II to Geometry sequence.

To observe course taking patterns prior to the CALPADS course enrollment data first available in AY 12-2013, I turn to California State Test (CST) files. CST-based measures of 8th grade Algebra participation are available from Spring 2003 through Spring 2013, when Course-specific end-of-course (EOC) exams were replaced by the SBAC. I use CST data only to extend the descriptive policy trend in Figure 1, which shows the rise and fall of 8th grade Algebra 1 enrollment across a longer time horizon. The CST and Course enrollment measures do not match exactly as CST records reflect Spring test-taking while Course enrollment is measured in the Fall (and may include duplicate student enrollments). I therefore do not use CST data to extend the analytic panel or construct analysis variables. CST Imputed Controls are examined as a sensitivity check in Table A5.

[6] Approximately one-third of district-years (903 of 2,767) have total 11th-grade math enrollment exceeding 100% of grade enrollment, reflecting students enrolled in multiple math courses simultaneously. The median excess is 6 percentage points. Extreme values occur in very small districts and do not meaningfully affect aggregate shares. I test directly whether the focal reforms changed doubling-up behavior and find no significant effects (Appendix Table A6).

Census Day Enrollment

I draw on census day school enrollment files from AY 2009-10 through AY 2022-23 also from the CDE, to scale course data as well as construct controls and weights reflecting district size. Changes in the course taking outcome variables are therefore interpretable as a percentage point change in the proportion of a grade level cohort enrolling in a specified course or course category. As with course enrollment data I sum CDE census day enrollment from the school-grade to district-grade unit. My main analysis excludes district-sponsored charter school counts in all years. This restriction preserves sample consistency across dependent variable categories and provides a clearer treatment-control contrast (i.e., charter schools did not always adopt the sequence type of the district sponsoring them).

Test Scores

11th grade district level SBAC scores for the California Assessment of Student Performance and Progress (CASPP) serve as my main measure of math proficiency. The SBAC assesses content aligned to the Common Core through Course 3 (i.e., Algebra II[7] or IM 3) but does not assess content beyond this level, such as precalculus or calculus. The test is required of all juniors regardless of course enrollment, eliminating selection into the tested sample. By design, the SBAC provides a common achievement benchmark across pathways. It was first administered to high school juniors in the spring of 2015 (i.e., the freshman class of fall 2012, the expected graduating class of 2016). I restrict the SBAC analysis to AY 2014–15 through AY 2018–19, excluding AY 2019–20 and later due to COVID-19 test cancellations, high opt-out rates, and the absence of parallel course data. This leaves between one and four years of pre-period test score data for pathway reform districts, depending on the year and grade level of initial treatment.

A note on data coverage and the CCSS implementation timeline

Before turning to the specifics of variable construction and sample definition, it is worth noting how the overlap of these data sources with California's phased CCSS implementation (Figure A1) shapes the analytic design. Although the state adopted the CCSS in August 2010, most districts did not implement aligned curricula and textbooks until around 2014-15. Integrated Math (IM) adoption was

[7] A handful of districts (N=4) primarily use an Algebra 1, Algebra 2, Geometry sequence. While I did re-code course-level enrollment data in these districts to ensure consistent course level classifications, my focus on 11th and 12th grades outcomes makes it irrelevant to my analysis.

heavily concentrated in that year, with 87 of 178 adopting districts first offering an integrated pathway in fall 2014. The 8th grade Algebra I (or Integrated Math I) rollback followed a parallel timeline with most of the decline occurring between 2013 and 2015 (in 8th grade timing). This maps to 11th grade outcomes in 2016 to 2018. The course enrollment panel (fall 2012-2018) and the SBAC panel (spring 2015-2019) therefore capture the principal wave of both policy changes, with two to four pre-treatment years for most treated districts. Pre-2012 course data would add limited analytic value: IM enrollment was literally zero before 2012 and eighth grade Math 1 shares were stable at approximately 60 to 65% across districts, with sharp divergence only beginning at CCSS implementation (Figure A1). Extending the panel forward is similarly constrained because the CDE paused course-level enrollment reporting after AY 2018-19 and COVID-19 disrupted both SBAC administration (spring 2020 was canceled statewide) and possibly the enrollment patterns that underpin the identification strategy. Post-2019 data therefore have limited value for the research design.

Measures

Integrated Math Adoption

I use the course enrollment data described above to define District IM adoption status by year and grade level. I construct a ratio of integrated course enrollment to total math enrollment in the course sequence (i.e., Algebra I, Geometry, Algebra II, Integrated Math I - III) in grade *g* and year *t*.

$$IM\_Share_{gt} = \frac{Integrated\ Math\ Enrollment_{gt}}{Integrated\ and\ Traditional\ Math\ Enrollment_{gt}} * 100\%$$

I define a district as having adopted an integrated math pathway when this ratio exceeds 50% among ninth graders. Consistent with district-level implementation, the distribution of this measure is sharply bimodal. Most districts are near 0% or 100% each year, with few in the middle (Figure 2B). The primary treatment variable captures cohort exposure: a two-year lag of ninth-grade IM status, which aligns the pathway a cohort entered in ninth grade with that cohort's eleventh-grade outcome. I use this measure to reflect the underlying theory of action of IM that a more coherent sequence of courses benefits students. Additionally, most IM districts appeared to introduce the reform in 9th grade and roll the sequence forward with that adopting grade rather than introduce it in all grades simultaneously. The time-varying binary treatment variable identifies the first year in which the lag measure crosses the

50% threshold and districts that never cross it are classified as never treated. Shifting this threshold does not substantively change any empirical results. I also construct a contemporaneous 11th grade treatment measure, which captures the share of current juniors enrolled in integrated course as a share of Math 1 through 3 enrollments, regardless of their initial 9th grade placement and use this measure as a robustness check. In the observed span, only one district adopted IM and then reverted to the traditional sequence. This district is counted as never treated in the main estimator but is allowed to switch when estimated using one of the concordance estimators.

8th Grade Algebra Rollback

The rollback treatment captures the withdrawal of early access to high school math in eighth grade. I define rollback using a level-based threshold rather than a change-based or continuous measure, for conceptual and methodological reasons. Conceptually the policy question concerns whether districts highly restricted or eliminated early access to high school math as a mainstream pathway. A district whose eighth-grade Math 1 share falls below 20% has throttled its acceleration program for an average student in a way one whose share declines from 65% to 45% has not. The 20% threshold captures this qualitative distinction.

Methodologically, continuous specifications of rollback intensity face several challenges in this setting. Continuous difference in differences methods remain at the econometric frontier (Callaway et al., 2025; de Chaisemartin & D'Haultfoeuille, 2024). The empirical tools for estimating causal effects of a continuous treatment with the same confidence as a binary treatment, particularly under staggered treatment timing, are not yet fully developed. I discuss these approaches and their limitations further in the analysis section. In this setting, the challenges are compounded by the fact that nearly all districts experienced some decline in eighth grade algebra enrollment during the CCSS transition leaving few stayer units for comparison. The level-based binary definition sidesteps these issues while preserving a clear policy interpretation: did the district nearly eliminate eight grade Algebra I as a mainstream pathway?

That said, the 20% level threshold is fundamentally arbitrary and cannot capture the full impacts of continuous rollback cleanly. I therefore examine alternative rollback thresholds and definitions as robustness checks. First, I identify *downstream* rollback, using the change in Math 1 enrollment in ninth grade as an intermediary treatment indicator. Here, a year-over-year increase of 20

more percentage points in 9th grade Math I share is defined as the onset of treatment, reflecting the arrival of rolled back cohorts in high school (Table 3, Panel C). This measure also has the benefit of allowing me to include high school only districts in the rollback analysis. Second, I use the pre-rollback baseline Algebra I share in 8th grade as a continuous dose measure, estimated via the continuous treatment difference in differences framework of Callaway, Goodman-Bacon, and Sant'Anna (2025; Table 3, Panel B). This allows me to account for the varying magnitude of the drop a district imposes when it crosses the 20% threshold. Third, I present results across a variety of alternative thresholds to approximate the dose-response plane. As I show in Section [N], point estimates are substantively consistent across these specifications, clustering between -2 and -3 percentage points on cohort-tracked advanced course-taking.

Course taking outcomes

The primary course taking outcome is the share of students enrolled in "All Advanced Courses", which includes Course 4 (Precalculus Statistics) and Course 5 (Calculus), other advanced in the six-level hierarchy previously described and detailed in Table B1. I examine this measure at 11th grade, 12th grade, and as a cohort-tracked combined measure that sums a cohort's twelfth-grade enrollment in year *t* with the same cohort's 11th-grade enrollment in year *t-1*. The denominator is the corresponding sum of 11th- and 12th-grade total enrollment, so the measure captures the share of person-grade-years in advanced courses across both upper grades. This nets out timing shifts: a reform that delays students from 11th-grade precalculus to 12th-grade precalculus without reducing their ultimate attainment will show offsetting effects at each grade level but no change in the combined measure. As such, I use it as my primary course-taking outcome. I also decompose the advanced category into component outcomes: precalculus (AP and non-AP), calculus (AP and non-AP), statistics (AP and non-AP), AP/IB math, Course 3 (Algebra II/IM III). An “any math” extensive margin measure confirms that the effects reflect changes in course composition rather than aggregate math participation.

Test score measures

The primary test score outcome is the district level mean 11th grade SBAC Math Z score. I standardize district-level mean scale scores within-year using the statewide norms published by the CDE. A proficiency decomposition (share meeting or exceeding standard) appears in the appendix. I

also construct subgroup-specific z-scores by race/ethnicity category (URM = Black, Hispanic, Native American, Pacific Islander) again using the statewide mean and standard deviation. I standardize 11th grade ELA SBAC scores in the same way for use in specifications that account for district ELA growth.

Control variables

The main specification for IM adoption includes a single time-varying control, the two-year lag of the ninth-grade math 1 share. This share captures the share of a cohort's ninth graders enrolled in algebra 1 or integrated math 1 at the time of their initial high school placement. Because rollback and integrated math were often co-incident (see Figure A2), this measure aims to isolate the independent effect of IM adoption by controlling for cross-district differences in the incoming pipeline (i.e., the extent to which a district had rolled back eighth grade algebra access by the time the tested cohort entered high school). For the rollback specifications, I include a time-varying IM adoption status variable, as a covariate for the same reason to separate the effect of rollback from any concurrent shift in course sequencing. Alternative controls (i.e., a three-year lagged eighth-grade Math 1 share, average ninth-grade course level, share in advanced courses at 11th grade, and CST-imputed controls) are examined in sensitivity analyses (Table 4, Table A4).

Analytic Sample

My sample construction begins with the 416 public school districts that report non-zero non-charter high school enrollment during the period where course data are observed. State-sponsored charter schools, state special schools (e.g., "California School for the Deaf"), and county offices of education are excluded given their atypical student populations contexts. I also exclude course and enrollment counts at district-sponsored charter schools from totals to match district composition on key outcome variables (district SBAC scores exclude charters) and to capture a stricter district-level treatment-control contrast. This removes approximately 7% of students from the sample. In secondary analyses I retain course-taking and enrollment data including charter school enrollment.

From the 416 districts, I retain the 395 that report state test data in each year of the pre-COVID analytic window (AY 2014-15 through 2018-19). This 395-district sample underlies all descriptive analyses (RQ1) and longitudinal IM test-score analyses. The rollback analysis uses 297 districts – the 330 that serve 8th grade with meaningful, reliably reported enrollment, minus those with baseline

(2012) Math 1 / Algebra I enrollment at or below 20%. Conceptually, such districts cannot transition into rollback treatment and could introduce floor effects to estimates. This threshold is inherently arbitrary; supplementary analyses present extensive robustness checks across alternative treatment definitions, thresholds, and operationalizations, including continuous estimators (Tables 2 and 4), verifying that results are not an artifact of threshold choice. All SBAC analyses are restricted to the pre-COVID period through spring 2019. Course analyses end in fall 2018, the last year of publicly available enrollment data.

Table A1 presents baseline characteristics for IM adopters (N = 228) and non-adopters (N = 167). IM adopters are somewhat smaller (K-12 enrollment of 9,936 versus 13,885), more likely to be in towns (26.3% versus 16.9%), and serve a higher share of Hispanic students (51.3% versus 40.6%, $p<0.01$) and students in poverty (22.9% versus 18.4%, $p<0.01$). Baseline course-taking patterns are comparable: 8th-grade high school math shares are 63.6% for adopters and 59.7% for non-adopters ($p=0.36$), and 9th-grade Math 1 shares are similarly balanced. At the first SBAC administration in 2014, adopters had modestly lower mean math z-scores (−0.080 versus 0.055, $p<0.01$) and lower 11th-grade advanced math enrollment (24.9% versus 31.4%, $p<0.01$).

These baseline differences, particularly the slightly lower initial achievement and advanced course taking among adopters, are consistent with IM adoption being more common among districts that were not already high-achieving and underscore the importance of a research design that accounts for pre-existing differences between treatment and comparison districts. My empirical strategy does not require baseline balance but would be threatened if districts diverged in time-varying characteristics over the panel. I test this directly with auxiliary regressions on total enrollment, racial/ethnic composition, and grade-level enrollment, and find no evidence of differential trends (Table A2). Additional detail on cohort coverage by reform type is included in the technical appendix

Policy Window 2: Qualitative data and sample

The quantitative analyses described above focus on the first policy window, the CCSS adoption period of 2013 to 2018. To capture how analogous dynamics may be playing out during the current adoption window, I draw on qualitative data from two sources collected under the Getting Down to Facts III (GDTF3) project during the 2024-2025 school year as California districts navigated the

curriculum adoption period following the release of the 2023 California Mathematics Framework. First, 94 semi-structured interviews with superintendents and district leaders were conducted by a PACE research team (Gallagher et al., 2026; see Appendix A for sampling methodology) using a stratified random sample of Unified school districts by region and district size, oversampling mid-sized districts (~50% response rate). Post hoc t-tests on demographic and performance variables across 15 region-by-size strata confirm that the sample is broadly representative of the underlying population of California Unified districts (Gallagher et al., 2026). Second, 38 semi-structured interviews with school principals were conducted by the author's research team from a stratified random sample of LEAs by district type, covering 69 districts across 38 counties (17% principal response rate). Both sets of interviews were recorded, transcribed, and anonymized. IRB approval was obtained from Stanford University.

Empirical Strategy

My empirical strategy relies on differential district-level take-up of two pathway reforms during the Common Core transition: (1) the staggered adoption of integrated math pathways for high school, and (2) the staggered rollback of 8th grade Algebra I (or Integrated Math I) access. Both reforms were prompted by the same policy trigger (i.e., CCSS adoption beginning in 2013) but they are analytically separable. As documented in Figure A2, many districts that never adopted integrated math nevertheless eliminated or dramatically restricted eighth-grade algebra (57% of IM non-adopters rolled back), and some integrated math adopters maintained near-universal early access (18% of Unified district IM adopters did not roll back middle-school algebra access). Moreover, even among districts that experienced both reforms, the timing was often staggered: approximately 46% of dual-treatment districts experienced the two reforms in different student cohorts, providing within-district temporal variation that helps distinguish the two effects.

A natural concern is that these effects reflect a broader deprioritization of advanced math in adopting districts rather than the specific policy decisions I study. The cohort-tracked design also partially addresses this: outcomes are measured two to four years after treatment and follow a specific student cohort through the course sequence, linking observed changes to discrete cohort exposures rather than to contemporaneous district-level preferences. A generalized shift in district priorities

would produce contemporaneous changes in course offerings; the lagged, cohort-following patterns documented here are harder to generate from a generic preference story.

I employ strategies derived from a difference-in-differences framework; wherein post-adoption changes in student outcomes and reform districts (i.e., treated units) are compared to contemporaneous outcome changes in districts that had not yet adopted the reform (i.e., control units). To illustrate the framework, I present the canonical two-way fixed effects (TWFE) estimating equation below:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_{it} + X_{it}'\delta + \varepsilon_{it} \qquad (1)$$

where Y_{it} is the outcome (SBAC z-score or course enrollment share) for district *I* in year *t*, α_i and γ_t are district and year fixed effects, D_{it} is the treatment indicator (IM adoption or post-rollback status), and X_{it}' is a vector of time-varying controls including a lagged measure of the other reform. The coefficient β captures the average treatment effect under the assumption of constant effects across cohorts and time.

This design rests on the identifying assumption that the difference in average outcomes across the treatment and control would remain constant in the absence of treatment. Because the counterfactual cannot be directly observed, this assumption is not explicitly testable. I investigate whether it may be supported empirically through event study specifications that test for differential pre-trends and I probe its sensitivity using the methods proposed by Rambachan and Roth (2023). The dynamic event study generalization replaces the single treatment indicator with a set of indicators for time relative to adoption:

$$Y_{it} = \alpha_i + \gamma_t + \Sigma_{\{k \neq -1\}\beta_k} \cdot 1\left[t - G_i = k\right] + X_{it}'\delta + \varepsilon_{it} \quad (2)$$

where G_i is district *i's* adoption year, *k* indexes event time relative to treatment, and k = -1 is the omitted reference period. Pre-treatment coefficients (β_k for k < -1) test for differential trends prior to adoption; their joint significance under the null H_0: $\beta_k = 0$ for all k < -1 is assessed via an omnibus Wald test. Post-treatment coefficients (β_k for k ≥ 0) trace out the dynamic treatment path, allowing effects to emerge or accumulate over time.

A second feature of this natural experiment that merits careful attention is the staggered timing of policy adoption, which causes the composition of the treatment and control groups to fluctuate annually. Recent econometric research has shown that static two-way fixed effects (TWFE) — which generalizes the canonical 2×2 DiD to multiple periods and groups — can produce biased estimates in this setting (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun & Abraham, 2021). The bias arises because staggered timing induces implicit comparisons between earlier- and later-treated units that generate invalid estimates when treatment effects are heterogeneous across units or time periods. Baker et al. (2022) demonstrate that these biases are empirically consequential in applied settings. I therefore implement heterogeneity-robust estimators, synthesized in a review of the emerging "new DiD" literature by Roth et al. (2023), that have been designed to recover valid estimates under staggered adoption.

Callaway and Sant'Anna's (2021) CSDID is my primary estimator. It avoids these "forbidden comparisons" by construction. I also report conventional TWFE estimates alongside three additional heterogeneity robust estimators as a concordance check on key outcomes. The Borusyak, Jaravel, and Spiess (2024) imputation estimator (BJS) recovers treatment effects by imputing counterfactual outcomes from untreated observations. The de Chaisemartin and D'Haultfoeuille (2024) estimator (dCdH) additionally permits treatment switching. For rollback analyses, I also report estimates from the Callaway, Goodman-Bacon, and Sant'Anna (2025) continuous treatment DiD (CGBS), which allows the differing baselines from which rollback districts reduced algebra take-up to be better accounted for. Agreement across estimators that make different assumptions strengthens their interpretation.

Primary Strategy: Callaway and Sant'Anna (2021)

My primary estimator is the group time average treatment effect on the treated (ATT) framework of Callaway and Sant'Anna (2021), implemented using the CSDID package in Stata (Rios-Avila et al., 2024). The core intuition is straightforward. Rather than pooling all treated units into a single regression as TWFE does, this estimator computes a separate treatment effect for each adoption cohort (defined by the year a treatment first becomes treated) in each post-treatment period. Formally, the estimator recovers a separate ATT for each period cell:

$$ATT(g,t) = E[Y_{t(1)} - Y_{t(0)} | G_i = g] \quad (3)$$

where ATT(g,t) is the average treatment effect for districts that first adopted in year g, evaluated at calendar year t. These group-time effects are aggregated into a summary ATT using the simple (equally weighted) scheme, which gives each group-time cell equal weight:

$$ATT = \Sigma_g \Sigma_t w(g,t) \cdot ATT(g,t). \quad (4)$$

Because each group-time ATT is estimated using only not-yet-treated districts as the comparison group, the estimator never makes the "forbidden comparison" that generates bias in TWFE. It never uses an already-treated district to estimate what would have happened to a later-treated district absent treatment. The not-yet-treated comparison group is larger than the never-treated alternative, improving precision in a panel with 395 districts and 5-7 time periods. It also relies on a weaker pre-trend assumption: trends need to be parallel only between each treated cohort and the contemporaneous not yet treated pool rather than between each cohort and a fixed set of never treated units over the full panel.

For both reforms, I restrict the treatment group variable to adoption cohorts observed by 2018 (i.e., the final panel year for course outcomes), excluding late-adopting districts whose treatment onset falls at or beyond the panel endpoint. These late adopters contribute no post-treatment observations but do contribute pre-treatment group-time cells to the omnibus pre-trend test, where their small group sizes and imprecise estimates can inflate the test statistic and generate spurious pre-trend rejections. Imposing this ceiling ensures that all treated cohorts contribute at least one post-treatment observation and that the pre-trend test reflects the identification-relevant portion of the data. In practice, this exclusion affects between 10 and 45 districts depending on the specification and lag structure, with larger exclusions for the combined cohort-tracked measures that require an additional year of lag. Point estimates are unchanged.

The group-time ATTs are aggregated using the simple (equally weighted) scheme, which gives each adoption cohort equal weight regardless of size. I report dynamic event study estimates that plot treatment effects by time relative to adoption and assess pre-trends using the omnibus Wald test, which jointly tests whether all pre-treatment event study coefficients equal zero. Inference is based on analytical standard errors derived from the influence function approach of Callaway and Sant'Anna (2021), clustered at the district level to account for within-district serial correlation and outcomes across time periods (Bertrand et al., 2004). The influence function approach yields pointwise

confidence intervals for each group-time ATT and for the aggregated parameters; simultaneous confidence bands for the event study path are computed using the multiplier bootstrap. I do not apply weights for district size in my main specification but test sensitivity to enrollment weighting.

Concordance Estimators

To assess sensitivity to estimator choice, I report results from three additional DiD estimators alongside CSDID. Each estimator addresses the staggered-timing problem through a different computational approach. BJS imputes counterfactual outcomes from untreated observations rather than computing group-time contrasts directly. dCdH additionally permits treatment switching, confirming results are not driven by the absorbing-treatment assumption. TWFE is reported for transparency; its concordance with the heterogeneity-robust estimators suggests that negative weighting bias is small in this setting, consistent with the compressed adoption timeline. Agreement across all four estimators strengthens the interpretation.

Addressing key identification concerns

Beyond the standard requirements of the DiD framework — namely, the parallel trends assumption and the complications introduced by staggered adoption timing — three empirical considerations structure the analysis that follows. First, the simultaneous implementation of the Common Core State Standards means that any improvements in math achievement among IM-adopting districts could reflect broader CCSS-driven gains rather than the integrated pathway itself. I address this through a triple-difference design that nets out district-level trends common to both math and ELA. Second, despite the analytical separability documented above, the frequent co-occurrence of the two reforms within districts means that isolating the independent contribution of each requires a multi-pronged identification strategy: single-treatment estimation with the other reform as a time-varying control, a joint TWFE specification that tests for interaction effects, and subgroup analyses restricted to districts that experienced only one reform. Third, the continuous nature of rollback (i.e., districts reduced eighth-grade algebra access by varying degrees) means that any binary treatment

definition involves an arbitrary threshold choice, which I address through dose-response analyses and a continuous treatment DiD estimator.

Simultaneity of CCSS implementation

To address the concern that IM effects on test scores may reflect broader CCSS implementation rather than the integrated pathway itself I implement two complementary strategies. The primary approach is a triple-difference (DDD) design that stacks math and ELA SBAC scores within each district-year, treating ELA as a within-district control subject that is exposed to the same CCSS-driven shocks but should not respond to math-specific pathway reform:

$$Y_{ist} = \alpha_{is} + \gamma_{it} + \lambda_{st} + \beta \cdot (D_{it} \times Math_{s}) + \varepsilon_{ist} \quad (5)$$

where *s* indexes subject (math or ELA), and the specification includes district×subject fixed effects (α_{is}), district×year fixed effects (γ_{it}), and subject×year fixed effects (λ_{st}). The coefficient β captures the effect of IM adoption on math achievement net of any district-level trend common to both subjects. If IM operates through math-specific channels (i.e., integrated instruction rather than general CCSS improvement), β should be positive even after absorbing the shared component. As a complementary check within the CSDID framework, I also estimate effects on the within-district math-minus-ELA z-score difference as the outcome (Table 1, Panel B), which nets out the shared CCSS component while preserving the heterogeneity-robust properties of the primary estimator.

Co-occurrence of IM and rollback

Because the two reforms frequently co-occurred (i.e., half of dual adoption Unified districts, Figure A2), I take several steps to confirm that the estimated effects are attributable to each reform individually rather than their joint adoption. I employ three complementary approaches. First, each CSDID specification includes a lagged measure of the other reform as a time-varying covariate. For the IM analysis, this is a lagged measure of ninth-grade Math 1 share, which absorbs simultaneous changes to middle school course assignment. For the rollback analysis, this is a lagged measure of IM adoption status. In both cases, the inclusion of this cross-reform control has negligible effects on the point estimates but its inclusion is conceptually important for isolating each reform's contribution.

Second, I estimate a joint TWFE specification that includes both reforms and their interaction simultaneously:

$$Y_{it} = \alpha_i + \gamma_t + \beta 1 \cdot IM_{it} + \beta 2 \cdot RB_{it} + \beta 3 \cdot (IM \times RB)_{it} + \varepsilon_{it} \quad (6)$$

If the two reforms operate through fully independent channels, the interaction term β3 should be approximately zero. A significant interaction would suggest that the reforms modify each other's effects, with a positive β3 indicating that IM partially offsets rollback-associated declines in districts experiencing both reforms. The CSDID framework does not currently permit simultaneous estimation of interacted treatments, but close agreement between CSDID and TWFE results across specifications lends credibility to this approach. I also estimate the TWFE interaction model separately for single-adopter districts (those that adopted only one reform, plus never-treated districts) and dual-adopter districts (those that adopted both, plus never-treated), which tests whether the interaction pattern differs across these subsamples. Third, I estimate separate CSDID models for each treatment subgroup — IM-only districts (N = 83), rollback-only districts (N = 71), and dual-adopter districts (N = 142) — each compared against 94 never-treated districts. These subgroups provide clean identification free of co-occurrence contamination and serve as a benchmark against which the full-sample estimates can be compared.

This benchmark is material because the joint TWFE interaction is not cleanly identified in staggered two-treatment settings: TWFE coefficients are weighted averages of 2×2 comparisons whose weights can be negative (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020), and the problem may compound when two treatments and their interaction are estimated jointly. The CSDID subgroup estimator does not rely on the separable functional form and therefore provides an independent check on whether the TWFE interaction captures separable treatment effects rather than artifacts of that functional form.

Rollback measurement

For the rollback analysis, the binary treatment definition involves an arbitrary threshold choice. To assess whether the results are robust to this discretization, I complement the binary CSDID estimates with the continuous DiD estimator of Callaway et al. (2025; CGBS). Rather than classifying districts as "rolled back" or not at a single cutpoint, this estimator uses the full continuous variation in

baseline eighth-grade Math 1 shares. The dose variable is each district's 2012 Course 1 share, which is fixed at baseline and captures the extent of potential rollback exposure: districts with higher baseline shares had more early access to lose and, empirically, experienced larger declines.

However, this does not fully capture the *time-varying* continuous treatment that actually occurred. Dose does not change within districts over the panel. The nascent continuous DiD literature has developed estimators for time-varying treatment intensity (e.g., de Chaisemartin & D'Haultfoeuille, 2024), but these require a sufficient number of "stayers" whose treatment value remains stable across periods. In this setting, the near-universal nature of rollback among high-baseline districts leaves too few stayers for reliable estimation. The baseline-dose approach is therefore the best available strategy for exploiting continuous variation in rollback exposure, though it captures cross-district differences in initial vulnerability rather than within-district changes over time.

An additional interpretive constraint applies to both the binary and continuous specifications: because the comparison group consists of districts that retained 8th-grade Math 1 above the 20% threshold but in many cases reduced it substantially over the panel, these estimates identify the effect of aggressive restriction or elimination of 8th-grade algebra access relative to partial reduction, not relative to a status quo of unrestricted access. Between 2012 and 2018, treated districts reduced 8th-grade Math 1 enrollment from 64% to 10%, while comparison districts moved from 67% to 46%, a roughly 33 percentage point differential reduction in the first stage. This attenuates estimates relative to a hypothetical 0%-vs-100% access contrast and represents a design-level floor on the achievable effect size.

The identifying assumption is a generalization of the standard parallel trends assumption to the continuous setting: conditional on baseline treatment intensity, the trajectory of untreated potential outcomes would have been the same for districts at different points in the dose distribution. Callaway et al. (2025) call this the "standard" or "weak" parallel trends assumption, and it is the direct analog of the binary case — it identifies the average treatment effect for districts at each dose level relative to no treatment. A key distinction for interpretation: this assumption is sufficient for estimating the level effect of a given dose (e.g., the effect of the policy shift on districts that started with 60% Math 1 access), but comparing effects across dose levels to recover a causal dose-response function requires a stronger assumption that restricts treatment effect heterogeneity across dose groups. Because I cannot

empirically distinguish between these assumptions, I interpret the continuous estimates as level ATTs at observed doses and treat the monotonic pattern across binary threshold cutpoints (Table A5, Panel B) as suggestive of, but not definitive evidence for, a dose-response relationship. The current implementation of CGBS does not support time-varying covariates, so these estimates are unconditional. I also report a series of binary "dose-response" CSDID estimates using progressively more inclusive threshold definitions (below 1%, 10%, 20%, 30%, 40%, and 50% Math 1 share; Table 2 Panel C) and parallel treatment definitions using percentage-point change thresholds in 8th and 9th grade (Table 3 Panel D). Binary CSDID and continuous CGBS estimates are also reported in Table 3.

Heterogeneity

I also explore whether the effects of both reforms vary across student subgroups and at different points of the achievement and course taking distributions. These analyses are particularly policy-relevant given the equity motivations cited by many reformers. First, I estimate separate CSDID models for underrepresented minority (URM) and non-URM students, defined as Hispanic, Black, Native American, and Pacific Islander. For test score outcomes, subgroup SBAC z-scores are standardized to the overall statewide distribution. For course taking outcomes, I use race-specific enrollment counts from the CALPADS course data as denominators. The race-disaggregated models use the same treatment definitions and estimation procedures as the main specifications.

Second, I estimate the effects on the share of students at each SBAC proficiency threshold: "Exceeded Standard," "Met or Exceeded Standard," and "Nearly Met Standard." This tests whether average effects mask heterogeneous responses at different points of the achievement distribution. Consistent with Ho (2008), I treat the continuous measure as primary and relegate the proficiency decomposition to the appendix.

Third, I estimate effects on lower-level course taking — specifically, the share of students enrolled in Course 1 (Math 1 or Algebra 1) at 11th and 12th grade. This is motivated in part by the rationale given for decelerating middle school students: in the "Algebra for All" era, Algebra I repetition was highly prevalent, and rollback proponents argued that students not rushed through algebra would be more successful the first time, reducing the need for retaking in later grades. While rollback districts show less ninth-grade retaking (Figure B2), the question is whether this translates into fewer 11th and

12th graders still enrolled in Course 1. A reduction would indicate that the initial gains from deceleration persisted through the pipeline.

Additional sensitivity and robustness

I assess the sensitivity of my main estimates to potential violations of the parallel trends assumption using the HonestDiD framework of Rambachan and Roth (2023). This approach constructs confidence sets for the treatment effect under the assumption that the post-treatment violation of parallel trends is no larger than some multiple ($\bar{M}$) of the maximum pre-treatment trend differential. The breakdown value reports the smallest $\bar{M}$ at which the confidence set includes zero — a higher breakdown value indicates greater robustness.

The IM analysis includes a lagged measure of ninth-grade course placement as a time-varying covariate. To ensure results are not sensitive to this choice, I report specifications with alternative control variables including a Unified subsample with an additional 8th-grade Math 1 control and a specification that adds contemporaneous 11th-grade course composition controls. I verify that the IM treatment estimate is stable across these alternatives in Table A4. As a measurement sensitivity check, I restrict the sample to districts whose maximum observed IM share is either 0% or at or near 100%, eliminating 46 partial adopters. This produces a binary treatment comparison between full adopters (N = 204) and never-adopters (N = 145), for an analytic sample of 349 districts.

I conduct several additional checks detailed in the appendix: (1) sensitivity to enrollment weighting (Tables A4, A5); (2) adding a contemporaneous 11th-grade course composition control to the IM SBAC spec (Table A4); (3) reintroducing charter-school enrollments in the lagged 9th-grade Math 1 control to test whether charter variation in the covariate affects the IM estimate (Table A4); (4) testing whether capping the 8th-grade Math 1 share at 100% affects treatment assignment (Table A5); (5) dropping districts with evidence of 8th-grade enrollment underreporting (Table A5); (6) treating coursecode 2498 ("Other") as Math 1 to bound the effect of course misclassification (Table A5); (7) excluding the largest districts by enrollment as a generalizability check (Table A5); and (8) effects on total math enrollment and Course 1 enrollment, which test for course-coding artifacts (Table A6).

Results

I present results in four sections, organized by outcome. The first examines the effects of both reforms on standardized test scores, including evidence on whether the integrated math effect can be separated from general CCSS-era achievement trends. The second estimates the effects of both integrated math adoption and rollback on high school math course composition, including evidence on the separability of their course taking effects from one another. Together, these sections address RQ2 and RQ3. The third examines RQ4 and answers whether either reform altered racial disparities in advanced course access. The fourth considers RQ5 and draws on practitioner interviews and surveys conducted during the 2024–25 curriculum adoption cycle to describe the current state of these policies and to examine whether the structural dynamics documented quantitatively during the CCSS transition are recurring as districts make new pathway decisions under the California Mathematics Framework. I then present a consolidated discussion of robustness and sensitivity analyses.

Effects on Math Achievement

IM Adoption and Test Scores

Table 1 presents estimates of the association between integrated math adoption and 11th grade SBAC math achievement. Two identification challenges frame this analysis: IM adoption occurred simultaneously with broad CCSS implementation, making it difficult to separate any IM-specific learning gains from general improvement; and IM adoption frequently co-occurred with rollback, raising the question of whether observed effects reflect the integrated pathway or concurrent changes to the course pipeline.

The preferred CSDID estimate yields an effect of 0.037 standard deviations ($p < 0.01$; Table 1, Panel A), with concordance across TWFE (0.027 SD, $p = 0.01$) and BJS (0.032 SD, $p = 0.01$; see Figure 4). The effect is small in absolute terms, though equivalent to roughly three months of 11th-grade learning (Bloom et al., 2008). The formal pre-trend test does not reject parallel trends ($p = 0.118$), though the estimate is fragile to even moderate deviations (see HonestDiD discussion below). Notably, the estimate is unchanged when controlling for 11th-grade course composition, suggesting the effect does

not operate through course-taking patterns. The achievement effect attenuates moderately when weighted by district enrollment to 0.017 SD ($p < 0.10$; Table A4).

Furthermore, auxiliary CSDID regressions on time-varying district characteristics show no evidence of differential changes to enrollment composition or level (Table A2). This supports the validity of the quasi-experimental design and is itself substantively notable: on average, parents did not strategically respond to IM or rollback through district entry or exit decisions.

Simultaneity: Separating IM from General CCSS Trends

Because IM adoption coincided with broader CCSS implementation, the 0.037 SD estimate may partly reflect general improvement among enthusiastic adopters rather than the integrated pathway itself. Two approaches address this. The math-ELA z-score difference (0.014 SD) is positive but imprecise, with a borderline pre-trend test ($p = 0.062$; Table 1, Panel A). The stacked DDD specification, which nets out any district-level improvement common to both subjects, produces 0.005 SD, indistinguishable from zero. These estimates suggest that the IM-specific contribution to average math achievement is small, on the order of 0.005–0.015 SD, and not distinguishable from broader CCSS-era gains with the available precision. That said, the proficiency decomposition suggests the DDD null may partly reflect averaging across the distribution: the share of students meeting the proficiency standard shows a positive DDD estimate (1.01 pp, $p < 0.10$; Table A3, Panel D). IM may have shifted some students across that threshold even if the mean effect is indistinguishable from general CCSS gains.

Rollback and Test Scores

The 8th-grade rollback produced no detectable effect on SBAC math achievement in the pooled CSDID but two complementary specifications indicate a small math-specific decline the pooled estimate does not capture. The math−ELA difference CSDID yields −0.026 SD ($p < 0.05$), and the stacked DDD returns −0.010 SD in the same direction, though not distinguishable from zero. The shared sign across all three estimators is lightly suggestive of a small rollback-induced depression in math relative to ELA.

The binary CSDID estimate is null across additional thresholds tested, from 10% to 50% of baseline Math 1 share and the proficiency decomposition (Table A3) shows null effects at every achievement threshold (Exceeded Standard, Met or Exceeded Standard, and Nearly Met Standard).

Co-occurrence: Disentangling IM and Rollback Achievement Effects

The TWFE interaction model (Table 1, Panel B) provides a complementary decomposition. When both reforms are estimated simultaneously, the IM main effect is small and positive but not individually significant ($\beta_1 = 0.012$), while the rollback main effect is negative and marginally significant ($\beta_2 = -0.025$, $p < 0.10$). Restricting the sample to single-adopter districts (i.e., those that adopted IM but never rolled back, or vice versa) yields very similar coefficients ($\beta_1 = 0.014$, $\beta_2 = -0.029$), reassurance that the main-effect estimates are not being driven by the co-adopting subsample.

The positive and marginally significant interaction term ($\beta_3 = 0.031$, $p < 0.10$) suggests that IM may have been modestly protective against rollback-associated learning declines: in districts that did both reforms, the IM and interaction coefficients roughly offset the rollback main effect. This pattern is echoed in the imprecise CSDID subgroup estimates (Table 1, Panel C): IM-only districts show +0.027 SD, rollback-only districts −0.030 SD, and both-reform districts +0.027 SD, consistent with IM absorbing the rollback's downward pressure. This congruence between the TWFE decomposition and the design-based CSDID subgroup ATTs reinforces confidence in the former.

Taken together, the evidence suggests that IM adoption was not harmful to student learning and may have been slightly beneficial, but the IM-specific contribution is difficult to isolate. After accounting for both CCSS simultaneity and rollback co-occurrence, the estimated IM-specific effect is in the range of 0.005 to 0.015 SD — a magnitude that is difficult to distinguish from zero and not practically meaningful on its own. The more defensible interpretation is that the integrated pathway produced learning outcomes at least comparable to the traditional sequence, and that whatever small gains appear likely reflect some combination of IM-specific benefits and the broader instructional investments that accompanied CCSS adoption.

Effects on Course-Taking

IM Adoption and Course Composition

Figure 5 presents the effects of IM adoption on high school course composition. The cohort-tracked All Advanced estimate shows a decline of 2.07 percentage points ($p < 0.01$, pre-trend $p = 0.054$; Table 2, Panel A), concordant with the BJS estimate of −2.83 pp ($p < 0.01$). This decline is offset

by a corresponding increase in Course 3 (Algebra II / IM3) enrollment of 4.34 percentage points ($p < 0.01$; Table B3). The All Advanced decline is concentrated in precalculus (−2.27 pp, $p < 0.01$; Table B3). Calculus, statistics, and AP/IB enrollment are statistically indistinguishable from zero. This IM course-taking decline could reflect some combination of forces: residual confounding from concurrent rollback not fully absorbed by the binary control, a broader deceleration orientation among IM-adopting districts, or the lack of acceleration infrastructure (e.g., doubling up, summer school) around the newly introduced integrated curriculum. The separability analysis below preserves the IM point estimate (−2.12 pp in the single-adopter subgroup) but the smaller subsample loses the precision needed to distinguish it from zero.

However, IM course-taking effects are entirely distinct from patterns downstream of rollback in several ways. IM reduces Course 1 (Algebra I / IM1) enrollment among upperclassmen: −2.53 pp in 11th grade, −1.20 pp in 12th grade, and −1.80 pp in the combined cohort-tracked measure (all $p < 0.01$; Table A6, Panel C). This is consistent with IM supporting a broader set of students through the core sequence and reducing the share still completing introductory coursework late in high school. Simultaneously, 12th-grade total math enrollment rises by 2.51 percentage points among IM cohorts ($p < 0.05$; Table A6, Panel A), indicating that more students remain enrolled in math through their senior year. This increase is not driven by course repetition, as Course 1 enrollment in 12th grade also declines, nor by double-enrollment, as the share of students exceeding 100% math enrollment is unaffected (Table A6, Panel B). The pattern is consistent with the integrated sequence encouraging more students to take a fourth year of math through continued progression rather than remediation. Such continuity in math programming may be beneficial for students who enroll in postsecondary math (Wainstein et al., 2023).

Rollback and Course Composition

The effect of reducing middle school algebra access to below 20% impacts advanced course-taking at the end of the pipeline, consistent with 8th-grade algebra access operating as a mechanical supply constraint rather than a broad preference shift. The cohort-tracked All Advanced measure declines by 2.74 percentage points in rollback districts ($p < 0.01$, pre-trend $p = 0.320$; Table 2, Panel A), offset by a mirrored 2.55 pp increase in Course 3 (Algebra II / IM3) enrollment as displaced students complete the standard sequence. Within the advanced basket, the decline lands on calculus

(−1.23 pp, $p < 0.01$; Table 3, Panel A), roughly 14% of the 8.7-point baseline, and precalculus (−1.10 pp, $p < 0.05$). Statistics and AP/IB effects are statistically indistinguishable from zero. Notably, the null on statistics rules out a pure calculus-to-statistics substitution within the advanced basket. If districts that restricted 8th-grade algebra had merely shifted advanced pathways toward statistics, we would expect a positive statistics ATT offsetting the calculus decline; we observe neither. The 12th-grade calculus estimate is larger (−1.90 pp, $p < 0.01$; Table A5, Panel A), but I rest identification on the cohort-tracked measure, which integrates across both grades and has clean pre-trends.

Four specifications converge on the aggregate estimate. The binary CSDID ATT of −2.74 pp is closely matched by the continuous-treatment CGBS estimate (−2.76 pp, $p < 0.01$; Table 3, Panel B), and dose-response across responsive-baseline thresholds is monotonic through the main specification: −2.74 pp at below 20%, −2.55 pp at below 30%, and −1.10 pp at below 40% (Table 3, Panel C). An alternative treatment definition based on 9th-grade Math 1 increases produces a larger effect of −4.00 pp ($p < 0.01$; Table 3, Panel D), consistent with the main estimates, though as a downstream measure I interpret these results cautiously.

As discussed in the methods, these estimates identify the effect of *near*-complete elimination of early access relative to partial reduction by comparison districts. The 2.74 pp decline in advanced course-taking and the mirrored 2.55 pp increase in Course 3 (Algebra II / IM3) together map a roughly 33 percentage point differential reduction in 8th-grade Math 1 enrollment (64% to 10% in treated districts vs. 67% to 46% in partially-treated comparison districts, 2012 to 2018) onto its later course-taking consequences. A 10 percentage point decline in 8th-grade algebra access therefore roughly corresponds to a 1 percentage point decline in advanced course-taking — a meaningful quantity, but also illustrative of the "leaky pipeline" to advanced math that preceded these reforms.

The null SBAC result alongside significant course-taking effects is consistent with what SBAC measures. The 11th-grade assessment tests grade-level content standards corresponding to Course 3 (Algebra II / Integrated Math III). Rollback shifts students from Course 4 (precalculus) to Course 3, but Course 3 students are still learning the content that SBAC assesses. Rollback also modestly reduces Course 1 (Algebra I / IM1) enrollment among upper-grade students (−1.22 pp combined, $p < 0.01$; Table A6, Panel C), consistent with the theory that deceleration reduces premature failure in 9th grade (Figure B2), but the effect is smaller than IM's (−1.80 pp). And, unlike IM, the share of students enrolled

in 12th grade declined by 2.4 percentage points following rollback. Students are not compensating for delayed entry to high school math by enrolling more frequently in a fourth year.

The most consequential sensitivity check for the rollback analysis is enrollment weighting. The unweighted CSDID estimate of −2.74 pp becomes small and statistically insignificant when districts are weighted by enrollment (Table 7; ATT = −1.08, pre-trend p = 0.019). Excluding the top 10% of districts by enrollment restores significance (−2.00 pp, $p < 0.01$), though the pre-trend test still fails ($p = 0.010$), indicating that very large districts drive the null. Dropping the top 20% returns a clean specification (−2.10 pp, $p < 0.01$, pre-trend $p = 0.275$), and dropping the top 30% returns a larger estimate still (−3.03 pp, $p < 0.01$). The most accurate reading is that rollback reduces advanced course-taking primarily in small-to-mid-size districts, which constitute most California districts though not the majority of students. Larger districts may better absorb the policy shift, perhaps because they can offer more alternative pathways. In Huffaker et al. (2025), for example, we find that SFUSD advanced math enrollment bounced back in the second post-reform year as alternative acceleration options (i.e., summer school, doubling up on Algebra I and Geometry, and an Algebra II and Precalculus compression course) expanded. The event studies for individual courses are too imprecise in the post-period to confirm or rule out recovery.

Separating IM and Rollback Course-Taking Effects

Because about four fifths of IM-adopting unified districts also rolled back 8th-grade algebra below 20%, a central question is whether these effects can be attributed to each reform individually. The TWFE interaction model (Table 2, Panel B) estimates both reforms simultaneously across all six course outcomes. For All Advanced, both IM ($\beta_1 = -3.50$, $p < 0.01$) and rollback ($\beta_2 = -5.01$, $p < 0.01$) show large, significant main effects, with a positive interaction ($\beta_3 = +3.33$, $p < 0.05$) indicating that IM partially offsets rollback-associated course-taking declines in dual-adopter districts. The same pattern appears across individual outcomes: rollback drives declines in precalculus ($\beta_2 = -1.31$, $p < 0.10$), calculus ($\beta_2 = -1.66$, $p < 0.01$), statistics ($\beta_2 = -0.94$, $p < 0.05$), and AP/IB ($\beta_2 = -1.36$, $p < 0.01$), while IM's independent effect on these courses is null or small. The calculus interaction is positive and marginally significant ($\beta_3 = +0.90$, $p < 0.10$), suggesting IM-adopting districts partially preserved calculus access even as rollback reduced it elsewhere. Course 3 results mirror the advanced-course pattern:

both IM ($\beta_1 = +3.76$, $p < 0.01$) and rollback ($\beta_2 = +5.17$, $p < 0.01$) shift students into Course 3, confirming the composition story from both directions.

Restricting to single-adopter districts provides the cleanest identification (Table 2, Panel C). Among single-adopter districts, IM-only shows no significant effect on All Advanced ($\beta_1 = -2.50$, ns), while rollback-only shows a large decline ($\beta_2 = -4.92$, $p < 0.01$). The dual-adopter subsample produces a rollback estimate that is too imprecise to separately identify (SE = 4.04), underscoring why the single-adopter comparison is essential. The CSDID subgroup analysis (Table 2, Panel D) reinforces the pattern: IM-only districts show a null effect (−2.12, ns), rollback-only districts show a −3.42 pp decline ($p < 0.01$), and dual-adopter districts show −3.77 pp ($p < 0.01$), with an implied interaction of +1.77 pp consistent with the significantly positive TWFE interaction. Dual-adopter districts carry an elevated but non-failing pre-trend ($p = 0.103$), warranting caution in reading the implied interaction. The CSDID recovers the same qualitative pattern as the TWFE, increasing confidence that the TWFE interaction captures separable treatment effects rather than functional-form artifacts.

The temporal co-occurrence heatmap (Figure A2) shows that 46% of dual-treatment districts adopted the two reforms in different years, providing within-district timing variation. Omitting the IM control from the rollback CSDID shifts the All Advanced estimate from −2.74 pp to −2.67 pp, a trivial difference that does not alter the qualitative conclusion (Table 7). Collectively, these analyses indicate that rollback is the primary driver of course-taking declines, with IM-only effects on advanced course composition statistically indistinguishable from zero. The two reforms are analytically separable in the single-adopter subsample, and IM appears to partially buffer rollback's course-taking declines in districts that adopt both.

Table A6 extends the analysis to other course outcomes. In 11th grade, neither reform affects total math enrollment (Panel A), and neither affects the share of students exceeding 100% math enrollment in either grade (Panel B), confirming that advanced-course effects reflect composition changes rather than aggregate math participation. In 12th grade, the two reforms diverge: IM increases total math enrollment by 2.51 pp ($p < 0.05$), while rollback reduces it by 2.36 pp ($p < 0.05$), consistent with IM supporting fourth-year math continuity and rollback pulling students out of the pipeline before 12th grade. IM also reduces combined Course 1 (Algebra I / IM1) enrollment by 1.80 pp ($p < 0.01$; Panel

C), with rollback producing a smaller decline of 1.22 pp ($p < 0.01$), consistent with both reforms reducing premature algebra repetition, but IM doing so more durably through 12th grade.

Additional Robustness and Sensitivity

HonestDiD

I assess sensitivity to parallel trends violations using Rambachan and Roth (2023; Table A7). For IM achievement, the original 95% CI narrowly excludes zero ([0.004, 0.049]) but includes zero at $\bar{M} = 0.5$ ([−0.013, 0.069]), indicating moderate sensitivity consistent with its small magnitude. For IM course-taking (combined All Advanced), the original CI is [−3.52, −1.28] and remains significant at $\bar{M} = 0.5$ ([−4.33, −0.52]), losing significance at $\bar{M} = 1.0$ ([−5.36, 0.46]). For rollback course-taking, the original CI is [−3.98, −1.40] and similarly remains significant at $\bar{M} = 0.5$ ([−4.72, −0.55]), losing significance at $\bar{M} = 1.0$ ([−5.87, 0.70]). Both course-taking effects are robust to pre-trend violations up to half the magnitude of the largest pre-period deviation, while the IM achievement effect is the most sensitive, consistent with its smaller magnitude.

Variable construction

As a measurement check, restricting to binary adopters (excluding districts with partial IM share) produces an identical IM SBAC point estimate (0.037 SD, Table A4). Adding an 11th-grade course-composition control yields a similar estimate (0.039 SD), suggesting the achievement effect does not operate primarily through course-composition shifts. Including charter students in the control also leaves the estimate unchanged (0.037 SD). Capping 8th-grade shares at 100% produces an identical rollback CSDID ATT (−2.74 pp; Table 7, variable construction block). Dropping districts whose 8th-grade math enrollment multipliers fell below 0.85 in 2017–18 (a proxy for underreporting) produces a similar estimate of −2.44 pp ($p < 0.01$). One measurement sensitivity is worth flagging: Spec C, which reclassifies 8th-grade "Other" coursecodes as Math 1 to bound the effect under maximum potential miscoding, attenuates the rollback estimate to −0.67 pp (ns). This indicates that the measurement of 8th-grade Math 1 enrollment is a meaningful source of uncertainty in absolute magnitude, though the direction and relative ordering of effects are consistent across specifications.

Comparison group and sample definition

Restricting to never-treated comparison districts produces estimates similar in direction to the headline results: IM SBAC 0.039 SD (Table A4); IM courses −2.12 pp in the IM-only subgroup (Table 2, Panel D), though this subgroup estimate is not statistically distinguishable from zero; rollback courses −3.14 pp (Table 7). Restricting the IM sample to unified districts produces a similar estimate (0.035 SD, Table A4); non-unified districts show a larger point estimate (0.053 SD) but fail parallel trends and are reported for transparency rather than identification. Three control-group variants for rollback produce estimates between −2.74 and −3.14 pp with clean pre-trends (Table 7, control-group composition block), including one that excludes late-rollback districts to address staggered adoption patterns (−3.10 pp, $p < 0.05$).

Effect Heterogeneity

To answer RQ4, I disaggregate the course-taking effects by group membership defined by historical representation in advanced courses (Table 4). This analysis reveals that both reforms narrow the racial gap in advanced course access, but through a leveling-down mechanism rather than equity gains in the conventional sense. IM adoption reduces advanced course-taking for non-URM students (−3.88 pp, $p < 0.01$; Table 4, Panel C), while the URM estimate is negative but not statistically distinguishable from zero (−1.38 pp, pre-trend $p = 0.022$, which fails parallel trends and warrants caution). The URM–non-URM gap narrows by 2.53 pp ($p < 0.01$, pre-trend $p = 0.462$; Table 4, Panel C), with clean pre-trends on the gap specification. The narrowing reflects leveling down: non-URM students lose more access, while URM students do not gain.

Rollback narrows the gap through the same mechanism. The combined gap narrows by 3.19 pp ($p < 0.01$, pre-trend $p = 0.231$; Table 4, Panel C). Non-URM combined All Advanced declines by 4.00 pp ($p < 0.01$) while the URM estimate is smaller and not statistically distinguishable from zero (−0.74 pp, ns), with clean pre-trends on both race-specific estimates. The TWFE interaction by race (Table 4, Panel D) reveals some nuance. Both IM and rollback main effects are large and significant for both groups: for URM students, β_1(IM) = −3.60 ($p < 0.01$) and β_2(RB) = −3.48 ($p < 0.05$); for non-URM students, β_1 = −5.01 ($p < 0.01$) and β_2 = −3.47 ($p < 0.05$). The IM × rollback interaction term is positive and significant for URM students (β_3 = +2.83, $p < 0.05$) but small and not significant for non-URM students (β_3 = +0.52, ns). The interaction gap (β_3 non-URM minus URM = −2.36, $p < 0.10$) indicates that IM buffers

rollback-associated declines more for URM students than for non-URM students in districts that adopt both reforms.

Practitioner Perspectives on Pathway Reform

The quantitative results suggest that both integrated math adoption and eighth-grade algebra rollback are associated with reductions in advanced course taking in 11th grade. Integrated math adoption coincided with a small improvement in test scores, though this finding is fragile and difficult to separate from broader CCSS-era trends; rollback produced null to slightly negative achievement effect. Both reforms narrow racial gaps in course access primarily by reducing access for non-URM students. To illuminate the mechanisms behind these aggregate patterns, I draw on interviews with 38 principals and 94 superintendents and district leaders across 69 California districts, conducted in fall 2025. Because the CALPADS panel ends in 2018, the interview and survey data also provide a window into whether the structural changes documented quantitatively have persisted into the present.

The superintendent interview sample was drawn via stratified random sampling from the universe of California Unified districts, stratified by district size and region, oversampling mid-sized districts, rather than by math pathway variables. Post hoc representativeness testing found no statistically significant differences between sample and non-sample districts on demographic and performance variables in 72 of 75 comparisons (Gallagher et al., 2026). The alignment between survey-reported pathway structures and 2018 administrative data (discussed below) provides some reassurance, but the sample was not designed to be representative on the dimensions most relevant to this study.

Table 5 compares pathway structures reported in the 2025 interviews and surveys against the 2018 CALPADS administrative data. In Panel A, 51% of principals report using integrated math, identical to the 51% IM adoption rate in the 2018 administrative data. The broader picture is one of continuity: the integrated math adoption that spread widely during the CCSS transition appears to have largely persisted rather than reversed in the seven years since the quantitative panel ended.

Table 1: Pathway Structures: 2025 Interview Data vs. 2018 Administrative Data

	2025 Interviews	2018 CALPADS
Panel A: High School Math Sequence (Principals, N=35)		
Integrated math	51%	55.9%
Traditional (AGA)	37%	44.1%
Both / Hybrid	11%	—
Panel B: Middle School Acceleration Access		
Students can take HS math before 9th	61% (principals) 62% (superintendents; Gallagher et al.)	—
Districts with ≥20% of 8th graders in Math 1	49% (supts; numeric, N=67)	34.2%
Among multi-track districts:		
Differentiation by 6th grade	43%	—
Differentiation in 7th grade	36%	—
Differentiation in 8th grade only	21%	—
Panel C: 8th Grade Algebra / Math 1 Enrollment (Superintendents, N=67; Gallagher et al.)		
Mean	19.8%	19.0%
Median	15.0%	11.8%
Panel D: Sequence × Acceleration Access		
	Principals (N=32)	*Districts (N=329)*
Integrated + Offers acceleration	10 (31%)	44 (13%)
Integrated + No acceleration	5 (16%)	137 (42%)
Traditional + Offers acceleration	7 (22%)	69 (21%)
Traditional + No acceleration	6 (19%)	79 (24%)

Notes: 2025 data from GDTF3 principal interviews (N=35 with classifiable sequence; N=32 with both sequence and Q5 acceleration in Panel D; 4 Both/Hybrid schools omitted from Panel D for clarity) and superintendent interviews (N=94; Gallagher et al., 2026, see Appendix A for sampling). 2018 CALPADS: 395 analytic-sample districts (Panel A); 330 districts serving 8th grade (Panels B–D; 329 in Panel D with non-missing IM share). Panel C compares supt-reported statistics across the 67 supts leading 8th-grade-serving districts who provided a numeric response against full-sample CALPADS statistics; matched supt–CALPADS identifiers are unavailable, so the comparison is between two samples drawn from the same population. CALPADS thresholds (Panels B, D): acceleration = ≥20% of 8th graders in Math 1 (matching the rollback identification cutoff used elsewhere in the paper); IM (Panel D) = ≥50% of 9th graders in integrated math. Panel B first row reports survey yes/no responses (Gallagher's coded composite of whether a district offers acceleration; no direct CALPADS analog because the underlying CALPADS variable is continuous). The second row binarizes supts' own numeric responses (8th-grade Math 1 share question, N=67) at the 20% threshold for apples-to-apples comparison with CALPADS.

Panel B addresses middle school acceleration. Among principals, 61% report that students can take high school math before ninth grade; among superintendents, 62% describe multiple math tracks. These figures lie between the 57% of Unified districts with at least 10% of eighth graders in Math 1 in the 2018 CALPADS data and the 66% share of surveyed principals reporting 8th grade Algebra I being offered at their school captured by the RAND American Math Educator Study in 2023 (Kaufman et al., 2024). Among multi-track districts, the point of divergence varies: 43% split students in sixth grade or earlier, 36% in seventh grade, and 21% only in eighth grade, revealing substantial heterogeneity in how and when districts sort students into differentiated pathways.

Panel C shows a similar pattern for eighth-grade algebra enrollment. Superintendent-reported shares in 2025 (mean 19.8%, median 15.0%) closely track the 2018 CALPADS values (mean 20.6%, median 14.3%). The modest decline in the mean and slight increase in the median are consistent with continued gradual rollback at the top of the distribution, but the overall picture is again one of continuity: the low rates of eighth-grade algebra documented in the quantitative analysis are not a transient feature of the early CCSS transition but a durable structural characteristic of California math education. Two themes from the interview data speak directly to the quantitative findings:

Curriculum adoption is a structural decision point for leaders

A consistent finding across both principal and superintendent interviews is that curriculum adoption cycles create occasions for revisiting the fundamental structure of math pathways. One principal whose school uses the integrated sequence explained: “This makes this year a math adoption year. So, we’ll be shopping curriculum, and at the same time, we will probably be making the decision as a team about whether we want to continue with an integrated sequence, or if we want to move back to traditional Algebra, Geometry, Algebra Two.” A superintendent described a similar dynamic: “We’re convening a committee to dig into textbook adoption as the guide, like a push and a pull in our district, is parents wanting high levels of math achievement. So graduating 8th grade with geometry, not just algebra.” Another described strategic paralysis: “We didn’t want to invest in textbook adoption and a change in program that would change again when we have to change Integrated Math into a four-year pathway.”

Districts feel pressure to offer early acceleration, are unsure how best to do so

The quantitative finding that rollback did narrow the URM/non-URM course taking gap, but primarily by reducing non-URM access rather than expanding URM access, highlights the challenge of trying to generate both high-achieving and equitable high school outcomes through middle school placement practices. The interview data suggest that leaders are aware that neither extreme that has characterized California's approach to offering high school math in eighth grade across recent decades – either universal acceleration or universal deceleration – works well. In some districts, access to accelerated pathways has historically operated through parent advocacy rather than systematic assessment. A principal described it plainly: “I think it’s the squeaky wheels that pursue it, so it’s not done basically like an assessment that says you’re ready for it. I think it’s opt-in, not based on any data.”

A superintendent offered a compensatory account for this behavior: “There were a lot of teacher kids getting into the classes because they were going to talk to the principals, and those kids shouldn’t have been in those classes. So, I tightened that up.”

Yet eliminating acceleration entirely creates its own problems, as the post-CCSS advanced math trends demonstrate. One principal warned: “If you think long term with that model, you’re going to push calculus out of high school. . . . 50% of the seats at your major universities are taken up by the private schools who push—those kids are still going to be doing calculus. So, if we take our public schools and push calculus out of it . . . we’re doing them a disservice.” Another principal described the competitive dimension more directly: “I think a lot of schools are losing enrollment to private schools because we’re not offering algebra.” Some districts are searching for a middle path. A middle school principal described using multiple measures: “We look at who are the 64 kids that are most deserving. And it’s not about which mom calls in July.” What emerges from the interviews is that *how* to make placement decisions in a way that responds to realistic pressures and reflects district values is an urgent question for many districts.

These practitioner accounts offer two insights that complement the quantitative analysis. First, they suggest that curriculum adoption functions as a policy window in which structural pathway decisions are bundled with instructional materials choices, potentially amplifying the effects of what appear to be pedagogical reforms. Second, they reveal that administrators are under significant pressure to provide acceleration pathways and are searching for ways to do so that are more systematic and equitable than the informal, advocacy-driven mechanisms that historically governed access. How to structure acceleration, rather than whether to offer it, is emerging as a central decision point for the current era of math pathway policies.

Discussion

The first contribution of this study is descriptive, documenting the scale of pathway reform during the Common Core transition. By 2018, more than half of the 395 districts in this panel had adopted an integrated math pathway for high school, and most unified districts had reduced Algebra I / Integrated Math I enrollment to below 20% of eighth graders. Nearly a third of unified districts essentially eliminated early access to high school math (below 1% of 8th graders). Given that the prior

"Algebra for All" era had failed to advance outcomes for lower-achieving students and had been associated with high Algebra I failure rates, some degree of deceleration was likely well-founded. However, such dramatic restrictions on access to early high school math have created national flashpoints out of districts like SFUSD (e.g., Karlamangla, 2024; Napolitano, 2024; Goldstein, 2026). Despite these changes to course sequencing and student acceleration having been entrenched for nearly a decade, little empirical evidence existed on their impacts to inform the debate that reignited over pathway policies following the release of the California Math Framework.

IM adoption coincided with a small achievement gain (0.04 SD) in the range of aggregate CCSS effects (Bleiberg, 2021), though how much is attributable to the integrated pathway versus broader CCSS implementation is unclear. IM districts had to invest in new materials, so the effect may partly capture taking curriculum reform seriously (Polikoff, 2015). The gain persists after controlling for course composition, suggesting it operates through instruction rather than pipeline position. IM also slowed mid-sequence course progression, reducing precalculus enrollment by 2.27 pp ($p < 0.01$). The TWFE interaction and subgroup decompositions suggest rollback accounts for most of the advanced-course decline when both reforms are estimated jointly, with IM-only effects not statistically distinguishable from zero. Notably, IM did not reduce calculus, statistics, or AP/IB enrollment, and was associated with more 12th-grade math enrollment (+2.51 pp, $p < 0.05$). This makes IM unlikely to be cost-effective as a standalone reform, but where adoption costs are already sunk, as during CCSS implementation or the current CMF materials window, an integrated approach is no worse, and may be modestly better for learning and fourth-year math continuity, than the traditional sequence.

The broad rollback of access to algebra in middle schools was an independent and unambiguous barrier to advanced course access: approximately 2.7 pp on the combined measure, with calculus declining by about 14% of baseline, and about 19% when the rollback effect is isolated from IM. No offsetting achievement gains were observed; if anything, rollback was associated with lower 12th-grade math enrollment (-2.36 pp, $p < 0.05$) and marginally lower math-minus-ELA proficiency. A small decline (1.2 pp) in upper-grade repetition of Math 1 was observed, though rollback did not extend the 12th-grade repetition reductions that IM produced. Together, the two reforms account for roughly 60% of the 4.4 pp decline in advanced course-taking between the 2015 and 2017 cohorts. Both reforms narrowed the URM–non-URM gap, but through leveling down (non-URM students lost

disproportionately more access) rather than expanding opportunity for underrepresented students. The evidence is sufficient to justify concern, but points away from adding frictions to acceleration and toward broadening the pipeline of students prepared for rigorous coursework in K–8. That structural pathway levers proved easier to pull than instructional improvement is, unfortunately, consistent with what we know about the relative tractability of these two reform strategies (Elmore, 2004).

Limitations and cautions

At the same time, several limitations of this report warrant consideration. First, the rollback of high school course offerings in middle school was so pervasive that capturing its full, continuous treatment impact poses a genuine methodological challenge that cannot be easily addressed using even cutting-edge econometric tools for panel data estimation. Additionally, rollback course taking estimates are sensitive to enrollment weighting, with effects concentrated among small-to-mid-size rather than very large districts, likely reflecting the greater institutional capacity of those districts to buffer deceleration through alternative pathways. The panel is short relative to treatment timing, limiting modeling of pre-trends and the ability to observe long-run outcomes (though this does allow the study to evade the empirical challenges imposed by COVID-19 trend breaks). The state assessment also does not capture learning specific to more advanced coursework.

Finally, my analyses of the rollback of "Algebra for All" do not capture the full universe of relevant students. Elementary (i.e., K-8) districts are not in CALPADS data and are therefore absent from both descriptive and confirmatory study of middle school algebra policy. Students who attended high school only districts are also excluded from most of the rollback analysis (i.e., excepting Table 3, Panel D which suggests their inclusion would attenuate course-taking results). A fuller accounting of the CCSS-era reform effects, particularly on students at the tails of the proficiency distribution, would benefit from capturing more of the population over a longer time-horizon. The latter may be possible when the CDE releases more contemporary course data.

State and district decisions about math pathways in the coming years will shape the academic trajectories of California students, and these findings suggest such decisions are indeed consequential. However, I urge nuance in interpreting and translating these results on several fronts:

***First, limits to explanatory reach and generalizability*:** I find that roughly two fifths of the 11th and 12th grade advanced math decline from 2012-2019 is *not* arithmetically accounted for by my analyses. Furthermore, advanced course-taking declines may reflect programmatic effects (e.g., hard to double up Integrated Math unlike Algebra I and Geometry; summer school and online course development may have lagged initial rollout) but could also indicate an attitudinal shift away from encouraging students into the most rigorous math courses available in 11th and 12th grade. The secular decline may also reflect this preference shift which is not necessarily policy malleable in the same way the mechanical relationship between middle school acceleration and high school course entry point is. Widespread critiques of calculus-as-capstone math programming (i.e., the pathway that requires that students negotiate a 5-course high school sequence) may influence contemporary patterns of course offering and selection (Shaughnessy, 2011b).

***Second, risk of overcorrection and ahistoricism*:** This paper does not argue that the turn away from middle school algebra was prompted ex nihilo by the CCSS. Rather, district reforms were primed by genuine shortcomings from the "Algebra for All" era. A significant body of high-quality research (e.g., Clotfelter et al., 2015; Domina et al., 2015; Lafortune, 2018; McEachin et al., 2020) suggests the premature acceleration of students who were neither prepared nor provided sufficient supports (e.g., double dose) to succeed in middle school Algebra served them poorly. The findings of this paper are best understood as complements to that literature, as well as to research that suggests acceleration of students with unfinished prerequisite learning may still be beneficial if aligned with comprehensive and often resource-intensive supports (e.g., Dee & Huffaker, 2026).

Third, putting scale in context: The statewide 40-plus percentage point decline in 8th-grade Algebra I enrollment reflected large within-district reductions (i.e., a median drop of 73 percentage points among rollback-treated districts) which translated into an approximately 3-pp decline in advanced course-taking and a 1-pp decline in calculus enrollment among those districts' students. These are proportionally meaningful effects relative to baseline. However, the high degree of attenuation from the first stage is a reminder that policy debates surrounding acceleration tend to focus intense attention on the presently small — though no more or less important than any other — group of students who ever reach calculus, AP math, and other advanced courses.

Implications for policy and practice

These findings have immediate relevance as California enters a new instructional materials adoption window under the 2023 California Mathematics Framework. The interview data collected for this study suggest that districts continue to navigate consequential pathway decisions amid competing pressures from parents, teachers, and the framework itself. Five lessons emerge.

1. District leaders are empowered during adoption windows.

The CCSS experience demonstrates that standards adoption can function as a catalyst for structural reform even when its influence on classroom instruction is modest. District administrators proved both motivated and well-positioned to reshape course sequencing and placement thresholds—decisions that directly shaped students' educational experiences at a scale and speed that instructional reforms rarely achieve (Loveless, 2021). For policymakers and advocates focused on the content of standards themselves, this is a reminder that the ancillary structural decisions standards adoption sets in motion may be at least as consequential.

2. Integrated math is a viable but not transformative option.

The integrated pathway produces learning outcomes at least comparable to the traditional sequence. When estimated simultaneously with rollback, IM's independent contribution to course-taking declines is smaller than the standalone estimate suggests and partially offsets rollback in dual-adopter districts, perhaps because more students in IM districts opt to take 12th grade math. For districts already investing in new materials and professional development during the current adoption cycle, the integrated sequence is a reasonable choice. The evidence does not support switching to an integrated pathway outside of a prescribed adoption window, where the costs of new materials, teacher training, and scope-and-sequence redesign would fall entirely on the reform itself.

3. Broad-based deceleration does not achieve its intended goals.

The rollback of middle school algebra was intended to pursue both equity and achievement: deeper engagement with foundational material could improve learning, and restricting acceleration could narrow gaps driven by differential access. The evidence suggests that neither materialized. Rollback produced no detectable improvement in test scores, only modestly reduced the prevalence of upper-grade repeaters, and narrowed racial gaps only by reducing opportunity at the top of the

distribution. As districts continue to grapple with acceleration policy, these findings argue against uniform deceleration as either an equity or achievement strategy. These course-taking effects were concentrated among the small-to-mid-size districts that constitute the majority of California districts; enrollment-weighted estimates suggest that larger districts were better able to absorb deceleration without measurable downstream effects on advanced course access, likely because they could offer alternative acceleration pathways.

4. Smarter, fairer acceleration requires moving beyond extremes.

Practitioners describe growing pressure to restore acceleration pathways, but also uncertainty about how to do so equitably. The interview data reveal that access has historically been mediated by parent advocacy (e.g., "the squeaky wheels,") rather than systematic assessment. The quantitative evidence suggests that the endpoints of the policy spectrum (i.e., everyone accelerated, or no one accelerated) both impose costs. A more promising path involves automatic, data-driven enrollment policies that identify students ready for advanced coursework through multiple measures rather than through the discretion of traditional gatekeepers.

5. Effective instruction on the road to algebra should be the priority.

More fundamentally, if the goal is to broaden the population of students prepared for advanced coursework while also strengthening the mathematical experiences of students who may never reach those classes, attention must be focused primarily on improving instructional quality in elementary and early middle school grades (e.g., consistent with Fuchs et al., 2021; National Mathematics Advisory Panel, 2008). The structural reforms studied here can rearrange when and with whom students encounter specific content, but they cannot alter the mathematical preparation prerequisite for success in more advanced coursework. This is a much harder problem than remixing course timing and sequencing, but it deserves centrality in policy discussions of K-12 math education.

References

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. *U.S. Department of Education.*

Allensworth, E. M., Cashdollar, S., & Gwynne, J. (2021). Improvements in math instruction and student achievement through professional learning around the Common Core State Standards in Chicago. *AERA Open*, 7(1), 1–18.

Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2), 370–395.

Becker, J. P., & Jacob, B. (1998). *Math War Developments in the United States* (Bulletin No. 44). The International Commission on Mathematical Instruction.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1), 249–275.

Binder, A. J. (2002). Contentious curricula: Afrocentrism and creationism in American public schools. Princeton University Press.

Bleiberg, J. (2021). Does the Common Core have a common effect? An exploration of effects on academically vulnerable students*. AERA Open*, 7, 233285842110107.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness,* 1(4), 289–328.

Boaler, J. (2016). Designing mathematics classes to promote equity and engagement. *Journal of Mathematical Behavior,* 41, 172–178.

Boaler, J., Schoenfeld, A., Daro, P., Harold Asturias, Callahan, P., & Foster, D. (2018, October 9). OPINION: How one city got math right. *The Hechinger Report*.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.

California State Board of Education. (2023). California Mathematics Framework. California Department of Education.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2025). Difference-in-differences with a continuous treatment (*NBER Working Paper No. 32117*). National Bureau of Economic Research.

Canaan, S., Mouganie, P., & Zhang, P. (2022). The long-run educational benefits of high-achieving classrooms (*IZA Discussion Paper No. 15039)*.

Carranza, R. (2015, March 10). Superintendent Richard Carranza's remarks on the secondary math sequence [Remarks to the SFUSD Board of Education]. San Francisco Unified School District.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2015). The aftermath of accelerating algebra: Evidence from district policy initiatives. *Journal of Human Resources*, 50(1), 159–188.

Cohen, D. K., & Ball, D. L. (1990). Policy and practice: An overview. *Educational Evaluation and Policy Analysis*, 12(3), 233–239.

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. National Governors Association Center for Best Practices & Council of Chief State School Officers.

Confrey, J., & Maloney, A. (2015). Engineering [for] effectiveness in mathematics education. In Large-scale studies in mathematics education (pp. 373–403). Springer.

Conger, D., Long, M. C., & Iatarola, P. (2009). Explaining race, poverty, and gender disparities in advanced course taking. *Journal of Policy Analysis and Management,* 28(4), 555–576.

Conrad, B. (2023, October 2). California's Math Misadventure Is About to Go National. *The Atlantic*.

de Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.

de Chaisemartin, C., & D'Haultfoeuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey (*arXiv:2112.04565*).

de Chaisemartin, C., & D'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics* (published online).

Dee, T. S., Huffaker, E., Phillips, C., & Sagara, E. (2022). The revealed preferences for school reopening. *American Educational Research Journal*, 63(2).

Dee, T. S., & Huffaker, E. (2026). Accelerating opportunity: The effects of instructionally supported detracking. *American Educational Research Journal*, 63(2).

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N., & Lewis, R. W. (2019). Beyond tracking and detracking. *Sociology of Education,* 92(3), 293–322.

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade algebra-for-all effort. Educational Evaluation and Policy Analysis, 37(3), 275–295.

Dougherty, S. M., Goodman, J. S., Hill, D. V., Litke, E. G., & Page, L. C. (2017). Objective course placement and college readiness: Evidence from targeted middle school math acceleration. Economics of Education Review, 58, 141–161.

Dykeman, K., Jackson, K., Kurlaender, M., Larson, B., & Reed, S. (2026). Does your math pathway make a difference? High school mathematics and college outcomes. *Getting Down to Facts III.*

Elmore, R. (2010). Getting to Scale with Good Educational Practice. Harvard Educational Review, 66(1), 1–27. https://doi.org/10.17763/haer.66.1.g73266758j348t33

Elmore, R. F. (2004). School reform from the inside out: Policy, practice, and performance. Harvard Education Press.

Fensterwald, J. (2014, August 12). New twist to old debate on accelerated math. *EdSource*.

Fensterwald, J. (2021, December). STEM professionals and professors slam California's proposed math framework. *EdSource*.

Fuchs, L. S., Bucka, N., Clarke, B., Dougherty, B., Jordan, N. C., Karp, K. S., Woodward, J., Jayanthi, M., Gersten, R., Newman-Gonchar, R., Schumacher, R., Haymond, K., Lyskawa, J., Keating, B.,

Morgan, S., & Jacobson, J. (2021). Assisting Students Struggling with Mathematics: Intervention in the Elementary Grades. What Works Clearinghouse.

Galley, M. (2004, May 5). California high schools lining up for algebra 1 waivers. *Education Week,* 23(34), 27.

Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis,* 22(3), 241–254.

Gao, N., & Lafortune, J. (2019). Common Core State Standards in California: Evaluating local implementation and student outcomes. *Public Policy Institute of California*.

Gallagher, H. A., Towne, L., Gomez, D. M., & Loeb, S. (2026). The California state role in supporting district capacity for TK-8 math improvement. *Getting Down to Facts III.*

Goldstein, D. (2026, March 24). San Francisco Killed 8th-Grade Algebra. Now It's Set to Come Back. *The New York Times*.

Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics*, 37(4), 1141–1182.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.

Han, S. W., Kang, C., Weis, L., & Dominguez, R. (2023). On track or off track? *Socius*, 9.

Harlow, J. (2015, October 27). California districts moving to new 'integrated' high school math pathway. *EdSource.*

Ho, A. D. (2008). The problem with "proficiency." *Educational Researcher*, 37(6), 351–360.

Hopkins, D. J. (2018). The increasingly United States: How and why American political behavior nationalized. *University of Chicago Press.*

Huffaker, E., Novicoff, S., & Dee, T. S. (2025). Ahead of the game? Course-taking patterns under a math pathways reform. *Educational Researcher*, 54(2), 91–102.

Kang, S. H. K. (2016). Spaced repetition promotes efficient and effective learning. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 12–19.

Karlamangla, S. (2024, May 24). How Algebra Became a Flashpoint in Schools. *The New York Times*.

Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. TBS The Book Service Ltd.

Koedel, C., Li, D., Polikoff, M. S., Hardaway, T., & Wrabel, S. L. (2017). Mathematics curriculum effects on student achievement in California. *AERA Open*, 3(1).

Kaufman, J. H., Covelli, L., & Holmes, P. (2024). Elementary and middle school opportunity structures for mathematics learning (Report No. RRA2836-2). RAND Corporation.

Lafortune, J. (2018). *Essays on the Distribution and Effectiveness of Educational Resources* [UC Berkeley]. https://escholarship.org/uc/item/8qh075sj

Loewus, L. (2015, February 24). Georgia to abandon mandate on teaching integrated math. *Education Week.*

Long, M. C., Conger, D., & Iatarola, P. (2012). Effects of high school course taking on secondary and postsecondary success. *American Educational Research Journal*, 49(2), 285–322.

Loveless, T. (2021). Between the state and the schoolhouse: Understanding the failure of Common Core. *Harvard Education Press*.

Loveless, T. (2023, May 16). California's New Math Framework Doesn't Add Up. *Education Next*.

Luna, J. (2023, April 13). Algebra for none: The effects of San Francisco's de-tracking reform. *Thomas B. Fordham Institute*: *Flypaper Blog*.

McEachin, A., Domina, T., & Penner, A. (2020). Heterogeneous effects of early algebra across California middle schools. *Journal of Policy Analysis and Management*, 39(3), 772–800.

Napolitano, J. (2024, March 4). San Fran ballot measure reflects 10-year battle to reinstate 8th-grade algebra. *The 74*.

National Mathematics Advisory Panel. (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. U.S. Department of Education.

Noh, J., & Webb, M. (2015). Teacher learning of subject matter knowledge through an educative curriculum. *The Journal of Educational Research*, 108(4), 292–305.

Oakes, J., Ormseth, T., Bell, R. M., & Camp, P. (1990). Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science. *RAND Corporation.*

Okano, K., & Tsuchiya, M. (1999). Education in contemporary Japan. Cambridge University Press.

Polikoff, M. S. (2012). Instructional alignment under No Child Left Behind. *American Journal of Education*, 118(3), 341–368.

Polikoff, M. S. (2015). How Well Aligned Are Textbooks to the Common Core Standards in Mathematics? *American Educational Research Journal*, *52*(6), 1185–1211.

Polikoff, M. S. (2017). Is Common Core "working"? And where does Common Core research go from here? *AERA Open*, 3(1), 1–6.

Pressman, J. L., & Wildavsky, A. (1973). Implementation: How great expectations in Washington are dashed in Oakland. *University of California Press.*

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.

Rosin, M., Barondess, H., & Leichty, J. (2009). Algebra policy in California: Great expectations and serious challenges. *EdSource*.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? *Journal of Econometrics*, 235(2), 2218–2244.

Schoenfeld, A. H. (2022, February 9). California's new math framework offers path to greater excellence and equity. *EdSource.*

Shaughnessy, J. M. (2011a). An opportune time to consider integrated mathematics. *NCTM.*

Shaughnessy, J. M. (2011b). Endless algebra—the deadly pathway from high school mathematics to college mathematics. *NCTM.*

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher,* 15(2).

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of Educational Research*, 81(4), 453–492.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.

Tarr, J. E., Grouws, D. A., Chavez, O., & Soria, V. M. (2013). The effects of content organization and curriculum implementation. *Journal for Research in Mathematics Education*, 44(4), 683–729.

Tintocalis, A. (2015, July 22). San Francisco middle schools no longer teaching "Algebra 1." KQED.

Tyack, D., & Cuban, L. (1995). *Tinkering toward Utopia: A Century of Public School Reform*. Harvard University Press.

Wainstein, L., College, R., Miller, C. E., Phillips, M., Yamashiro, K., & Melguizo, T. (n.d.). *Twelfth Grade Math and College Access*.

Will, M. (2014, November 10). In transition to Common Core, some high schools turn to "integrated" math. *Education Week.*

Xu, Z., & Cepa, K. (2018). Getting college-ready during state transition toward the Common Core State Standards. *Teachers College Record*, 120(6), 1–32.

Technical Appendix

Imputed Course Data

The imputed observations are Farmersville Unified (2016), Modoc Joint Unified (2014), Calaveras Unified (2014), San Pasqual Valley Unified (2014), Banning Unified (2017), Klamath-Trinity Joint Unified (2016), Patterson Joint Unified (2016), and Shoreline Unified (2018). In each case missing values were filled by linear interpolation (or by using the neighboring observation for terminal years). Exeter Unified and Lammersville Joint Unified are excluded from the analytic sample because their recent formation dates (2015 and 2011, respectively) leave early panel years structurally missing or cannot be legitimately imputed.

Analytic Sample Continued

IM cohort coverage

Because no district offered an integrated math pathway before 2012, there are no observable pre-data adopters. Of the 395 districts in the analytic panel (sample_full_expanded), 228 ever adopted IM at 9th grade. Of these, 210 have L2 cohort exposure years (adoption year + 2) that fall within the capped panel window (2015–2018) and are identified as treated. The remaining 18 districts adopted IM at 9th grade in 2017 or 2018, so their L2 cohort exposure years (2019 and 2020) extend beyond the 11th-grade outcome panel; they appear as never-treated by construction in both the SBAC and course analyses. Among the 210 identified IM adopters, 24 that adopted 9th-grade IM in 2013 (L2 cohort = 2015) are always treated in the SBAC panel. Their treatment turns on in the first year of SBAC administration, so they contribute to post-treatment ATTs but have no pre-treatment SBAC observations and cannot contribute to pre-trend tests. These districts are fully identified in the course enrollment panel with multiple pre-treatment years. Dropping these 24 districts yields modestly attenuated but substantively similar estimates.

Rollback cohort coverage

Of the 297 unified districts in the rollback analytic sample (sample_rb_expanded, baseline 8th-grade Math 1 share > 20%), 234 ever crossed below the 20% threshold during the panel. The remaining 63 districts either held their baseline 8th-grade Math 1 share above 20% throughout or only

partially reduced it. Among the 234 crossers, 22 crossed in 2016, 2017, or 2018, placing their downstream 11th-grade cohort outcome years (2019–2021) beyond the course panel; these districts are classified as never-treated by construction. As with the IM ceiling, this restriction reflects the limits of the observation window rather than a modeling choice.

After this restriction, the 11th-grade (L3-based) rollback analysis captures 212 of 234 districts that ever crossed the 20% threshold (90.6%). Treatment timing for the L3-based spec is distributed across three 8th-grade cohorts: 2013 (57 districts, L3 gvar = 2016), 2014 (113, L3 gvar = 2017), and 2015 (42, L3 gvar = 2018). The cohort-tracked combined spec uses L4 timing and identifies 170 districts, dropping the 42 districts that crossed in 2015 because their L4 gvar (2019) falls outside the panel. The ceiling cap also prevents small late-adopter cohorts from destabilizing the omnibus pre-trend test, as discussed in the methods section. For both reforms, I restrict the treatment group variable to adoption cohorts observed by 2018 (the final panel year for course outcomes), excluding late-adopting districts whose treatment onset falls at or beyond the panel endpoint. These late adopters contribute no post-treatment observations but do contribute pre-treatment group-time cells to the omnibus pre-trend test, where their small group sizes and imprecise estimates can inflate the test statistic and generate spurious rejections.

Panel truncation and cohort observation windows

Figure A1 maps the pipeline from treatment to observed outcomes for each adoption cohort. The earliest rollback cohort (2013 at 8th grade, 57 districts) is observed through 11th grade (2016–2018) and 12th grade (2017–2018) with multiple post-treatment years, providing the strongest identification. The largest cohort (2014 at 8th grade, 113 districts, 53% of the L3 sample and 66% of the L4 cohort-tracked sample) is observed at 11th grade in 2017 and 2018 (two post-treatment years) and at 12th grade in 2018 only (one post-treatment year). This truncation is a limitation of the panel: long-run effects, including any recovery of advanced course access by 12th grade, may not yet be visible for the majority of treated districts in the cohort-tracked combined spec. The mostly-null rollback estimates at 12th grade for non-calculus outcomes should therefore be interpreted cautiously. They may reflect genuine catch-up, but they may also reflect insufficient post-treatment exposure for the dominant cohort. The cohort-tracked combined measure partially addresses this by pooling 11th and 12th grade outcomes for the same cohort, but cannot speak to effects beyond the observed panel.