The Effects of High School Curriculum Standards Reform: Evidence from Texas*

Andrew Smith†

November 7, 2025

Click here for the latest version

Abstract

This paper combines design-based evidence with a structural model to examine how curriculum standards influence student choices and school incentives. I exploit a Texas policy that made a college-preparatory curriculum the default for incoming ninth graders. Using a two-way fixed effects design, I show that it improved high school and postsecondary outcomes for low-performing students. To uncover mechanisms and assess alternative designs, I estimate a dynamic schooling model of curriculum assignment, course-taking, and college enrollment, and use it to simulate three potential policies. First, requiring all students to begin in the college-preparatory track raises college attendance by 12 percentage points for low-performing students. Second, state education policies prioritizing college enrollment increase attendance from 48% to 54% but raise dropout by 3 percentage points. Third, policies emphasizing test performance narrow achievement gaps by 10% and reduce dropout to 6.5% at the cost of lower college attendance. These results highlight clear trade-offs between short-term achievement and long-term college access, underscoring the need to align school incentives with postsecondary goals.

*I am grateful to Chao Fu, Chris Taber, Jeff Smith, Naoki Aizawa, Jesse Gregory, and Matthew Wiswall for their invaluable advising and patient support. I am grateful for financial support from the UW-Madison Department of Economics, and the National Bureau of Economic Research. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas.

†Department of Economics, University of Wisconsin-Madison (email: jasmith52@wisc.edu).

1 Introduction

Prior research has shown that high school curriculum choices play a central role in shaping students' human capital, influencing both short- and long-term outcomes (Goldin and Katz, 2007; Hanushek and Woessmann, 2011). Yet, despite its importance, high school curriculum design remains relatively understudied by economists, even though it represents one of the most direct ways policymakers influence students' skill development and postsecondary opportunities. In practice, states adopt curriculum standards aimed at improving college readiness while enforcing accountability systems that measure school performance using short-term metrics such as standardized test scores. Schools respond strategically to these policies by reallocating resources or tracking students by ability to raise measured performance, even when such adjustments may reduce college preparation (Gamoran, 1992; Betts and Shkolnik, 2000; Macartney et al., 2021). These decisions affect students' course choices and peer environments, ultimately influencing educational outcomes. Although student choices and school incentives jointly determine secondary and postsecondary trajectories, much of the literature examines these effects separately.

In this paper, I examine how high school curriculum standards and their assignment affect students' course choices in high school and their postsecondary enrollment. I address two questions: (1) How does high school curriculum reform influence students' course choices and postsecondary enrollment? and (2) How do school-level tracking policies shape postsecondary access?

To answer these questions, I exploit a Texas policy[1] that restricted how schools could assign students to curriculum standards. Before the reform, students could be placed into either a high (college-preparatory) or low standard, with the latter barring immediate entry into a four-year college.[2] The policy shifted the default by requiring students to begin in the college-preparatory track,[3] thereby raising the level of coursework most students completed in high school. Leveraging this policy variation and rich administrative data from Texas, I combine design-based and structural modeling approaches to examine three mechanisms: how students adjust their course choices in response to curriculum standards, how schools' incentives shape assignment decisions, and how peer effects influence learning outcomes and college-going. Together, these elements provide a comprehensive analysis of how curriculum reforms affect both short-run academic investments and longer-run postsecondary outcomes.

To begin my analysis, I estimate a two-way fixed effects (TWFE) model to measure the impact of making the college-preparatory curriculum the default for students identified as "at-risk"

[1] TAC Chapter 74, Subchapter E.

[2] Students graduating under the low standard were eligible to enroll in a two-year institution but were typically required to first complete remedial coursework covering the college-preparatory courses not completed in high school.

[3] This policy did not guarantee completion of a college-prep curriculum, since students could still graduate with a non-college-prep diploma or drop out.

of dropping out prior to high school.[4] I find that raising default standards reduces dropout rates by about 3 percentage points (p.p.) and increases graduation with the most difficult diploma by roughly 8 p.p. among students considered at-risk prior to entering high school. For this group, achievement scores decline by 0.12 standard deviations (s.d.), but enrollment in Algebra II rises by 5 p.p., and postsecondary attendance increases by about 2 p.p. These results suggest that while higher standards may impose short-term costs in the form of lower test scores, they expand postsecondary access by ensuring that more students complete the coursework prerequisites for college.

These results illustrate a fundamental tension in curriculum policy: schools must balance the risk of over-assigning students to rigorous tracks against the risk of under-assigning those who could succeed in them. Assigning students to a track that is too demanding can lead to dropout, whereas keeping capable students in less demanding tracks limits their preparation for college-level work. My findings suggest that the latter, under-assignment, has larger consequences for students' long-run outcomes, motivating an analysis of how schools make these assignment decisions.

To understand the mechanisms underlying the policy's effects and to evaluate alternative designs, I develop and estimate a two-stage dynamic schooling model of decision making by both school administrators and students. In the first stage, I model the equilibrium of high school curriculum assignments, where schools allocate students to specific standards and students respond through their course enrollment decisions. Students then realize their high school achievement through standardized test scores and learn their graduation status. In the second stage, they decide whether to enroll in a postsecondary institution. To connect the model to the empirical evidence, I estimate it using indirect inference with importance sampling (Sauer and Taber, 2021), a simulation-based approach that recovers structural parameters by matching simulated data to key patterns in the Texas data, including my design-based estimates.

This model uniquely highlights the interdependence between schools' strategic behavior and students' educational trajectories. Schools act strategically by (1) assigning students to standards in ways that shape peer composition and instructional environments, and (2) balancing short-term accountability goals tied to standardized tests against longer-term objectives related to college enrollment. Student decisions depend on track assignment, ability, test performance, and peer characteristics, which interact to determine course-taking and postsecondary choices. Ability influences both assignment and achievement, while track placement affects the peers and instructional quality students experience, linking school incentives to individual outcomes.

With the estimated model, I simulate three policy counterfactuals to evaluate how changes in

[4] Under Texas law (TEC §29.081), an "at-risk" student is legally defined as any individual under age 21 who meets specific statutory criteria, most commonly related to poor academic performance, such as failing courses, standardized assessments, or being retained in grade, but may also include factors such as pregnancy, disciplinary removal, justice system involvement, homelessness, or foster care placement. The state requires schools to provide additional support to at-risk students to help them succeed.

curriculum standards and school incentives affect student outcomes. First, I simulate a scenario in which all students are initially assigned to the college-preparatory curriculum. This policy substantially benefits students who would have otherwise entered the lower standard by raising their postsecondary enrollment rates by about 12 p.p. For students who would have been assigned to the higher curriculum, high school dropout rates increase by 1 p.p., and postsecondary enrollment falls by 3 p.p. Second, I examine an accountability policy that directs schools to prioritize college enrollment over achievement scores. Postsecondary attendance rises from 48 to 54 percent, but the dropout rate increases by about 3 p.p., illustrating the trade-off between short-term completion and long-run attainment. Finally, I simulate an accountability policy that increases the value schools place on achievement in their accountability objectives. This approach improves short-term outcomes: standardized test performance increases by 0.06 standard deviations and the achievement gap[5] narrows by roughly 10 percent, while dropout declines from 9.5 to 6.5 percent. However, postsecondary enrollment falls from 48 to 45 percent, showing that greater emphasis on test performance can hinder college access. Overall, these counterfactuals show that curriculum standards can meaningfully expand opportunity, but whether high standards translate into both higher achievement and postsecondary access depends on how school incentives are designed and aligned with long-run goals.

This paper contributes to several strands of literature. First, it relates to research on curriculum standards and the role of high school rigor in promoting college readiness. Completing more demanding coursework in high school, whether through higher grading standards or expanded course offerings, is strongly associated with college enrollment, degree completion, and labor market success (Figlio and Lucas, 2004; Rose and Betts, 2004; Joensen and Nielsen, 2009). However, few studies have examined how differences in curriculum standards, whether assigned or chosen, shape students' course-taking decisions and access to postsecondary education. I contribute to this literature by providing new evidence on how default curriculum standards affect both course-taking and college enrollment, combining design-based evidence with a structural model that captures how schools and students jointly respond to curriculum policy.

Second, this paper adds to a large literature on student tracking and peer effects. A long line of research examines how grouping students by ability affects achievement and longer-term outcomes, though evidence remains mixed. Some studies find little effect of tracking in high school (Betts and Shkolnik, 2000; Figlio and Page, 2000), while others suggest that tracking benefits lower-ability students at the expense of higher-ability peers (Argys et al., 1996; Fu and Mehta, 2018). Recent work in European settings explores how students' self-selection into academic and vocational tracks influences educational attainment (Cockx et al., 2019; De Groote, 2023;

[5] In what follows, the achievement gap refers to the difference in the state's terminal standardized mathematics test scores between students identified as at risk of dropping out prior to high school and those who were not.

Humphries et al., 2023). I extend this literature by modeling both sides of the assignment process: schools strategically allocate students to tracks, and students respond through course-taking and postsecondary choices. My model allows track assignment to influence both high school behavior and preferences for two- versus four-year colleges. I further incorporate nonlinear peer effects that shape instructional rigor and college attendance, providing a richer characterization of how peer composition interacts with ability and curriculum assignment.

Finally, this paper contributes to research on school accountability and the incentives created by performance-based systems. Prior work shows that accountability policies tied to standardized test outcomes can meaningfully alter school behavior. Some studies find that accountability can raise achievement (Deming and Figlio, 2016), while others document strategic responses and unintended consequences (Figlio, 2006; Koretz and Barron, 1998; Macartney, 2016). Additional research emphasizes the limits of such systems when reward functions are poorly aligned with broader educational goals (Muriel and Smith, 2011; Figlio and Loeb, 2011). I contribute to this literature by using a structural model to evaluate how alternative accountability designs shape both school behavior and student outcomes. Through counterfactual simulations, I quantify the trade-offs between policies that emphasize short-term test-based accountability and those that reward postsecondary enrollment. This approach provides new evidence on how accountability incentives interact with curriculum policy to influence educational sorting, achievement, and long-run college access.

The remainder of the paper is organized as follows. Section 2 introduces the institutional setting. Section 3 describes the data and provides descriptive evidence. Section 4 outlines the design-based strategy and presents the corresponding results. Section 5 develops the model, and Section 6 discusses model estimation, identification, results, and model fit. Section 7 provides some comparative statics from the estimated model. Section 8 presents the results from the counterfactual simulations, and Section 9 concludes.

2 Institutional Setting

2.1 The Policy Landscape of High School Diplomas

In the United States, not all high school diplomas signal the same level of academic preparation, either across schools or even within the same school. By 2014, states offered a total of 93 distinct diploma types, reflecting substantial differences in the expectations placed on students (Achieve, 2015). Figure 1 summarizes this landscape by showing which states require coursework that meets "college- and career-ready" (CCR) standards.[6]

While all diplomas certify high school completion, only some certify readiness for college-level work. In 26 states, multiple diploma options are offered, typically including at least one that meets CCR standards and another that is less rigorous. Among these, 14 designate the college-ready pathway as the default, requiring students to opt out (or fail out). In contrast, 20 states offer only diplomas that fall short of CCR standards, and four states plus the District of Columbia offer only the CCR option. Although students everywhere can choose to take more advanced coursework, in many states their diplomas do not signal this readiness to colleges or employers.

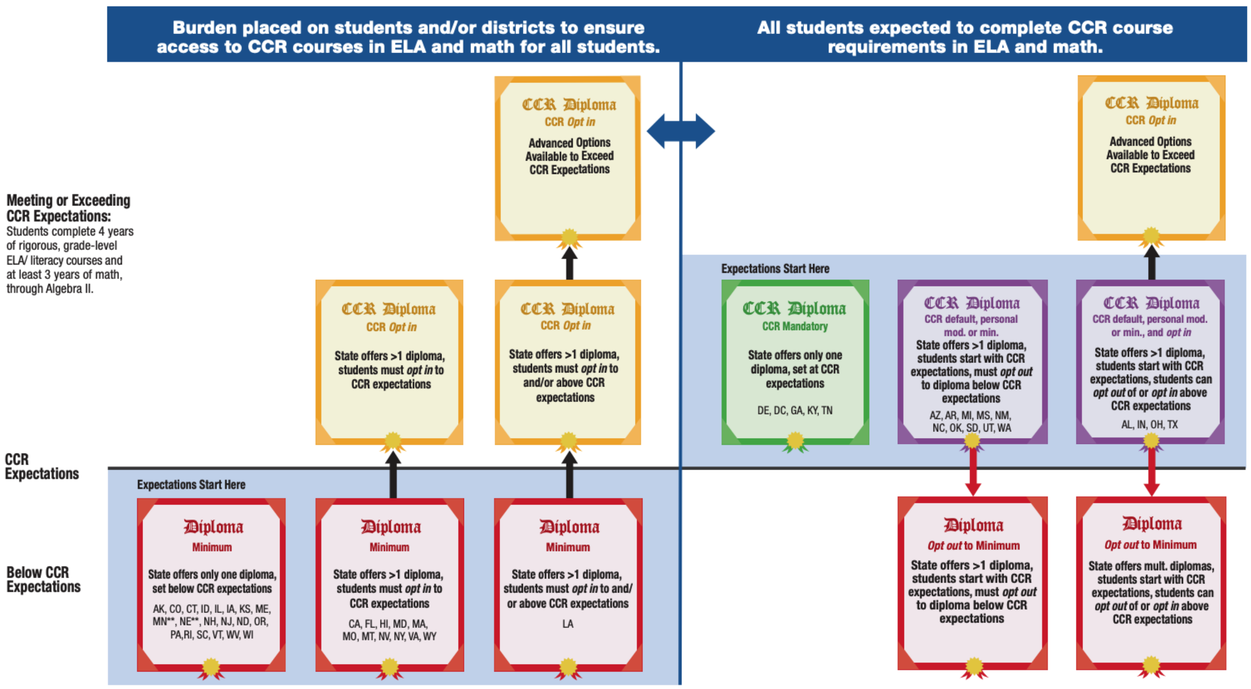


Figure 1: Graduation Expectations, by State (Source: Achieve (2015))

Texas provides an early and influential example of this policy design. Before most states differentiated diploma options, Texas adopted multiple graduation pathways and, in 2005, became the first to make the college-ready curriculum the default for all students (Achieve, 2015). Aligning graduation requirements with admission standards at public four-year universities created a natural setting to study how default curriculum standards influence students' course-taking and postsecondary access, which I discuss next in Section 2.2.

[6]Achieve, Inc. defines a CCR diploma as one requiring at least three years of mathematics (including Algebra II) and four years of college-preparatory English.

2.2 The Texas Curriculum Reform Under *Closing the Gaps*

In October 2000, the Texas Higher Education Coordinating Board (THECB) adopted *Closing the Gaps by 2015* (CTG), a statewide higher education plan aimed at closing educational disparities within Texas and between Texas and other states. The initiative targeted four objectives: student participation, student success, excellence, and research (THECB, 2000). One of CTG's key priorities was to increase college readiness among Texas high school graduates, prompting changes to high school graduation policy.

To advance these goals, the Texas Legislature enacted TAC Chapter 74, Subchapter D, requiring the Texas Education Agency (TEA) to update the state's high school graduation requirements (TAC, 2000). Beginning with the 2001–02 cohort of entering ninth graders, schools implemented three distinct diploma programs: the Minimum High School Program (MHSP, or "low standard"), the Recommended High School Program (RHSP), and the Distinguished Achievement Plan (DAP). The RHSP and DAP curricula aligned with the course requirements for admission to Texas public four-year universities, while the MHSP did not. In what follows, I refer to the RHSP and DAP jointly as the "high standard."[7]

Table 1: Minimum Course Requirements, by Curriculum Standard

Subjects	Low Standard	High Standard
English	English I English II English III Additional Course	English I English II English III English IV
Math	Algebra I Geometry Additional Course	Algebra I Geometry Algebra II
Science	Biology IPC	Biology Additional Course Additional Course
Social Studies	2.5 Credits	3.5 Credits
Electives	7.5 Credits	5.5 Credits
Foreign Language	0 Credits	2 Credits

Table 1 summarizes the key differences between the two standards. On the extensive margin, the high standard requires Algebra II and English IV in addition to the baseline courses in the low standard. On the intensive margin, students in the high standard exchange four of the seven and a half elective credits permitted under the low standard for one additional upper-level science,

[7]The DAP represents an enhanced version of the RHSP that requires the same coursework plus additional distinguished elements. For analytical purposes, both are considered part of the high track.

one upper-level social studies course, and a two-year foreign language sequence. Although both standards cover similar subjects, their timing differs: students on the low standard typically take remedial courses in ninth grade to prepare for Algebra I and English I.

Prior to the 2004-05 school year, schools were permitted to sort students through these curriculum tracks. Tracking through differentiating curriculum is a common practice in Texas public schools, and Antonovics et al. (2022) find that this form of within-school tracking far exceeds any between-school sorting on ability. To determine a student's curriculum assignment, schools assigned students to curriculum standards based on their performance prior to high school.[8]

Assignment to the low standard effectively limited a student's graduation pathway. Students placed in the low standard rarely graduated under the high standard, since "catching up" would require completing additional coursework outside the regular academic year. Moreover, the implications of track assignment were often not salient to students or families.[9] Within a high-stakes accountability regime, school administrators thus had incentives to assign lower-achieving students to the low standard to protect overall performance metrics.

The THECB and TEA later recognized that such early assignment practices could undermine CTG's objectives, especially for students at risk of dropping out. To address these concerns, the Texas Legislature repealed and replaced TAC Chapter 74, Subchapter D, with Subchapter E (TAC, 2003). The amendment made the high-standard curriculum the default for all entering ninth graders beginning in the 2004–05 school year. Students who might ultimately graduate under the low standard were now required to begin in the higher track.

Subchapter E also created clear conditions for transitioning to the low standard. A student could enroll in the low track only if they met one of the following criteria: (i) were at least sixteen years old, (ii) had completed two core credits in each subject area, or (iii) had failed promotion to grade ten one or more times. In such cases, the student, a parent or guardian, and a counselor were all required to consent to the change.[10]

2.3 Accountability Pressures and Strategic Assignment

Texas evaluates school performance through an accountability system administered by the Texas Education Agency (TEA), which rates schools and districts based on student achievement, growth, and other performance indicators (TEA, 2022). Although this system is intended to improve student outcomes, it also shapes the incentives schools face when making curriculum assignment

[8]A current Texas district administrator noted that Algebra II completion was often viewed as the key hurdle to obtaining the high-standard diploma.

[9]Giustinelli and Pavoni (2017) document similar issues in Italy, where students demonstrated limited awareness of available tracks and their consequences, particularly among disadvantaged groups.

[10]Students typically fail promotion by performing poorly during the school year and on their end-of-course exams, though local district policies may vary.

decisions. When performance metrics emphasize short-term outcomes such as test scores or graduation rates, schools may assign students strategically across curriculum tracks to raise measured performance, thereby altering peer composition and the level of teaching instruction in ways that can improve school-level achievement in the short term. However, these same adjustments may limit students' exposure to rigorous coursework and reduce long-run college preparation. Understanding these forces is essential for interpreting how curriculum standards operate in practice, both in Texas and in other states that employ similar accountability systems.

A substantial literature documents how accountability policies encourage schools to focus narrowly on short-term test metrics rather than broader measures of learning or long-term success (Figlio and Loeb, 2011; Deming and Figlio, 2016). Koretz and Barron (1998) show that schools in Kentucky engaged in "teaching to the test," while Figlio and Getzler (2006) find that Florida schools strategically reclassified students as disabled to exclude them from aggregate performance calculations.[11] These examples illustrate that when accountability systems place high stakes on measured outcomes, schools often prioritize improving reported performance rather than enhancing underlying student learning. Even when schools do not act strategically, the complexity of accountability formulas can limit their ability to respond effectively, as administrators may struggle to identify which student metrics most affect ratings.

In the Texas context, the design of the accountability system magnified these challenges. The TEA's accountability framework relied on intricate formulas that combined test scores, graduation rates, and subgroup performance measures. The 2006 Accountability Manual alone spanned over 200 pages, making an ex-ante computation of potential ratings virtually impossible for school administrators. As Muriel and Smith (2011) note, such complexity makes it difficult for schools to translate policy incentives into targeted actions. In practice, schools often focused on simpler metrics, such as average exit-level math scores, the share of students passing state exams, or dropout rates, to guide their efforts.[12]

These accountability pressures interacted directly with the curriculum assignment policies described in Section 2.2. Before the 2004–05 cohort of ninth graders, Texas high schools had full discretion to assign students to either the low or high curriculum standards. The 2005 reform restricted this by making the high-standard curriculum the default for all students. If peer effects played an important role in determining accountability ratings, such as through aggregate test performance, schools would have been more strategic in their standards assignments before this policy change. Focusing on short-term metrics like test scores may have improved accountability ratings in the moment but potentially harmed students' long-run outcomes. A system that instead

[11] Figlio (2006) further shows that some Florida schools imposed harsher disciplinary measures on low-performing students during testing periods.

[12] Based on discussions with district administrators in Texas, who emphasized that schools typically prioritized metrics like exit-level math performance, the share of students meeting proficiency standards, and dropout rates.

emphasizes longer-term indicators, such as postsecondary enrollment, could yield very different results. Understanding these behavioral responses is central to interpreting the effects of the reform analyzed in the sections that follow.

3 Data

3.1 Data Description and Sample

For this paper, I use administrative data from the Texas Schools Project (TSP), maintained by the Education Research Center (ERC) at the University of Texas at Dallas. The TSP links student-level records from the Texas Education Agency (TEA), the Texas Higher Education Coordinating Board (THECB), the National Student Clearinghouse (NSC), and the Texas Workforce Commission (TWC), enabling comprehensive tracking of students from K–12 into postsecondary education and the labor market.

My analytic sample consists of repeated cross-sections of students entering ninth grade between the 2002–03 and 2006–07 academic years. Throughout the paper, I refer to each cohort of ninth-grade students by the spring semester of their freshman year (e.g., the 2002–03 cohort of ninth grade students is referred to as the "2003 cohort"). I include all students who either graduated or dropped out of the Texas public school system.

From TEA enrollment records, I collect a range of background characteristics, including gender, race/ethnicity, free or reduced-price lunch (FRPL) status, and at-risk of dropping out status. I also compile all available standardized test scores from grades 3 through 11. During this period, students were assessed using the TAAS and TAKS exams, and I standardize these scores by grade and test year.

I merge each student to their semester-level course enrollment records for both middle and high school. In addition to enrollment, I observe whether a student passed each course in a given semester. These records cover core subjects (math, science, English, and social studies) as well as foreign languages, and career and technical education (CTE).

To capture postsecondary outcomes, I link student records to the THECB and NSC data, which together document enrollment, persistence, and graduation at public and private 2-year and 4-year institutions. From these datasets, I construct measures of college enrollment, degree completion, and declared major.

In addition to student-level records, I incorporate school- and district-level data from two sources: the Common Core of Data (CCD) and the Academic Excellence Indicator System (AEIS). The CCD provides annual demographic and institutional characteristics, including FRPL shares, Title I status, and racial/ethnic composition, at both the school and district levels. The AEIS includes a broad set of school performance measures, such as accountability ratings, exit-level TAKS and TAAS passing rates, commended performance indicators, and graduation rates.

Finally, I link teacher employment records to each school-year observation, allowing me to construct measures of average teacher characteristics, including educational attainment, tenure, experience, and salary.

3.2 Summary Statistics

Summary statistics by high school exit type are presented in Table 2. Each cell reports the mean (with standard deviation in parentheses) of a given characteristic, conditional on whether a student graduated from the low track, the high track, or dropped out of high school.

Table 2 reveals that students who drop out or graduate under the low standard are more likely to be from underserved groups, economically disadvantaged students, and those identified as at-risk prior to entering high school. These groups are also disproportionately male. Notably, being classified as at-risk or economically disadvantaged in middle school are among the strongest predictors of high school dropout.

Table 2: Summary Statistics, by High School Exit Type

	Total	Dropout	Low Standard	High Standard
Hispanic	0.402	0.599	0.355	0.387
	(0.490)	(0.490)	(0.479)	(0.487)
Black	0.133	0.205	0.160	0.124
	(0.340)	(0.404)	(0.367)	(0.329)
Asian	0.036	0.009	0.015	0.041
	(0.187)	(0.095)	(0.124)	(0.199)
Male	0.489	0.517	0.618	0.472
	(0.500)	(0.500)	(0.486)	(0.499)
At-Risk	0.371	0.800	0.603	0.304
	(0.483)	(0.400)	(0.489)	(0.460)
Economically Disadvantaged	0.421	0.701	0.452	0.390
	(0.494)	(0.458)	(0.498)	(0.488)
Post-Secondary Enrollment	0.541	0.080	0.306	0.612
	(0.498)	(0.272)	(0.461)	(0.487)
Math Std. Score (8th Grade)	-0.000	-0.984	-0.574	0.159
	(1.000)	(0.705)	(0.735)	(0.971)
N	756,985	61,110	69,310	626,565

Dropouts perform significantly worse on the state's eighth-grade standardized math assessment, scoring nearly a full standard deviation below the mean, while students graduating under the high standard score substantially above average. Although overall postsecondary enrollment in the sample is relatively high (54%), enrollment rates vary sharply by exit type: only 8% of dropouts *eventually* pursue postsecondary education, compared to 61% of high standard graduates.[13] This pattern underscores the strong correlation between curriculum rigor and postsecondary access.

Finally, while eighth-grade math scores are highly predictive of exit type, the standard deviations within each group indicate substantial overlap in the distribution of prior achievement. This suggests that high school sorting is not determined solely by academic readiness and that other factors–such as school assignment practices or student preferences–also contribute meaningfully to exit outcomes.

3.3 Descriptive Evidence

In this section, I provide preliminary evidence showing how the change in assigment rules impacted student high school and postsecondary outcomes.

3.3.1 Cohort Exit Rates

Figure 2 presents the distribution of high school exit types by entering ninth-grade cohort. Panel (a) shows trends for all students, while panel (b) restricts the sample to those flagged as at-risk of dropping out prior to high school.

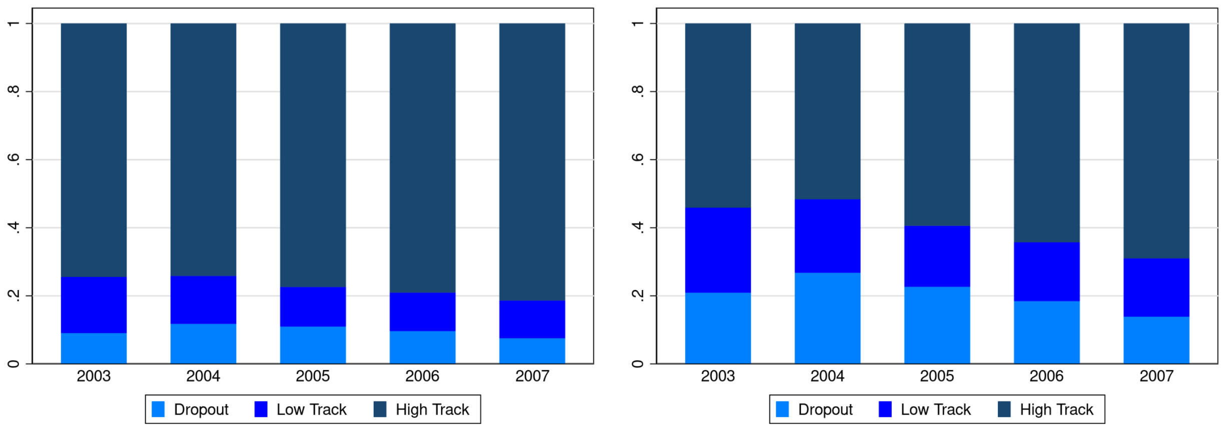
In panel (a), there is a clear upward trend in the share of students graduating under the high standard, with a discontinuous jump between the 2004 and 2005 cohorts. This timing coincides with the implementation of the assignment rule change, suggesting that the reform effectively restricted schools' ability to assign students to the low standard. As a result, a greater share of students completed the high standard curriculum. At the same time, the share of students completing the low standard modestly declines beginning in 2005, supporting the interpretation that the policy shifted students upward in the curriculum distribution.

Although the policy was implemented at the school level, it was designed to encourage students at the lower end of the ability distribution to complete college-preparatory coursework. A key subgroup that schools and states use to identify such students consists of those classified as at risk of dropping out. Panel (b) shows the distribution of exit types for this subgroup, and the effects of the policy are even more pronounced than in the full sample. The decline in low-standard assignments is steeper, and the increase in high-standard completions is more substantial. This

[13] In the context of the model, college enrollment occurs immediately following high school. While some dropouts may later obtain a diploma and pursue postsecondary education, this pathway is excluded from the model.

pattern suggests that the compulsory high-standard assignment policy was successful in reaching its intended target population.

Notably, dropout rates decline over time for both the full sample and the at-risk subgroup. This pattern suggests that the increase in high standard assignments did not come at the expense of higher dropout rates. For at-risk students in particular, the combination of declining low standard assignments and stable dropout rates implies that many were able to meet more rigorous graduation requirements when schools were required to assign them to a higher standard.



(a) All Students (b) At-Risk Students

Figure 2: Share of Exit Types by Ninth Grade Cohort

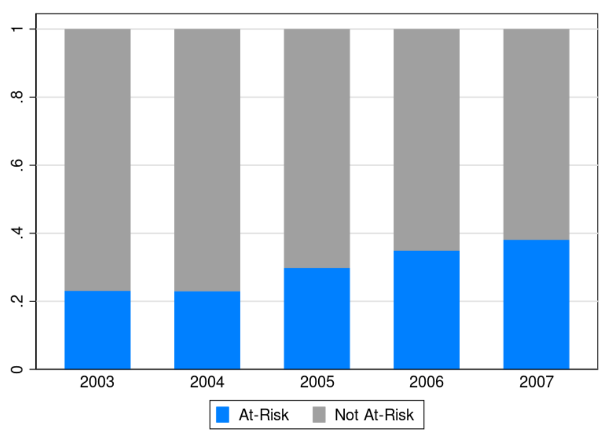


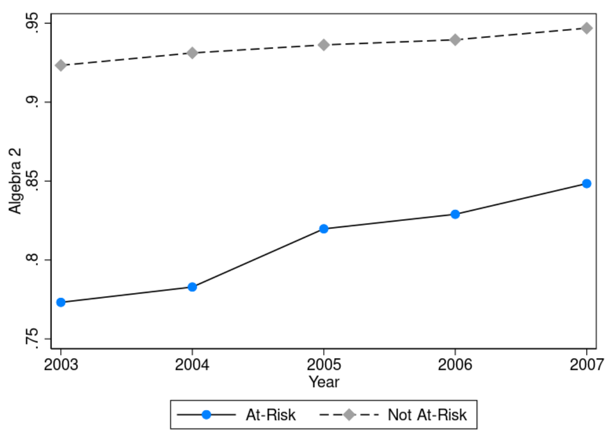
Figure 3: Share At-Risk and Non-At-Risk Students among High Standards Graduates

Figure 3 provides an alternative perspective on these dynamics by focusing specifically on the composition of students graduating under the high standard. Among early cohorts, only about one in five high-standard graduates were classified as at-risk. Following the 2005 policy change, however, the share of at-risk students among high-standard graduates rises sharply, reaching nearly forty percent by the 2007 cohort. This shift highlights that the reform did not simply expand the number of high-standard diplomas, but also broadened access to more rigorous coursework for students who historically would have been less likely to complete it.

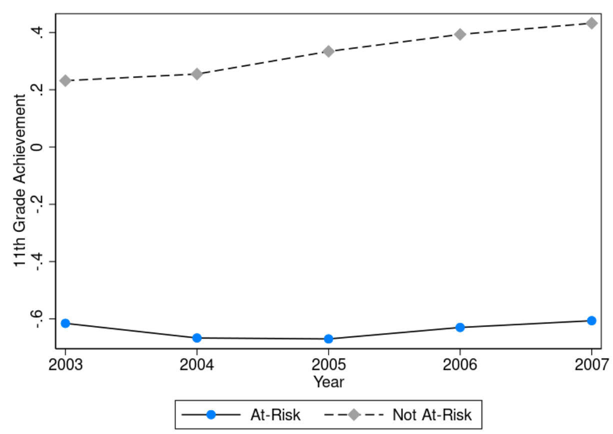
Taken together, these cohort trends provide descriptive support for the model's key institutional mechanism: that schools alter curriculum assignments in response to policy incentives, and that these decisions can meaningfully affect student trajectories.

3.3.2 Trends in Outcomes Over Time

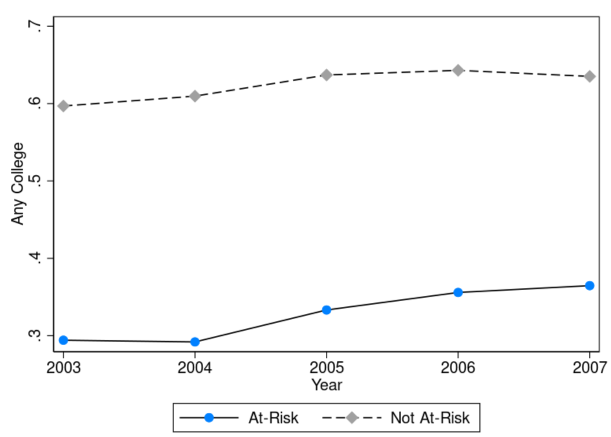
To complement the descriptive evidence on exit types, I next examine trends in coursework, achievement, and postsecondary outcomes across cohorts.



(a) Algebra 2 Completion



(b) 11th Grade Achievement



(c) Postsecondary Attendance

Figure 4: Student Outcomes, conditional on At-Risk Status

Figure 4a shows a sharp increase in Algebra II completion among at-risk students beginning with the 2005 cohort, while not-at-risk students remain consistently high. This suggests that the assignment policy directly altered the course-taking of students who otherwise would have been placed on the low standard.

Despite these shifts in coursework, Figure 4b indicates that large achievement gaps between at-risk and not-at-risk students persist. At-risk students make only modest progress, while their peers continue to improve steadily, pointing to limits on how much curriculum assignment alone can close underlying disparities.

The longer-run effects are reflected in Figure 4c, which shows steady gains in overall college attendance among at-risk students following the policy change. Although differences by risk status remain, the narrowing gap suggests that the reform expanded postsecondary access for students who were previously less likely to enroll.

These descriptive trends illustrate how the reform reshaped student pathways: increasing exposure to advanced coursework, shifting achievement, and broadening college access. These patterns motivate both the reduced-form and structural analyses that follow, which provide sharper evidence on impacts and mechanisms.

4 Design-Based Analysis

To quantify the descriptive evidence presented in Section 3, I estimate a series of regressions that assess the policy's effects on student outcomes. I begin with a standard two-way fixed effects (TWFE) model, where students the state has identified of being at-risk of dropping out prior to high school are considered treated, and their non-at-risk counterparts serve as controls. For robustness, I then extend this approach using the synthetic TWFE estimator of Arkhangelsky et al. (2021), which reweights units and time periods to match pre-treatment trends between treated and comparison groups, thereby relaxing the traditional parallel trends assumption.

4.1 Two-way Fixed Effects

The goal of this section is to estimate the impact of assigning the high-standard curriculum by default to students who were low-performing prior to high school. To do so, I begin with a standard two-way fixed effects (TWFE) specification, estimated at the school-by-risk-status level. Let s index schools, $r \in \{0,1\}$ indicate at-risk status, and t index academic cohorts. Define At Risk_{sr} as an indicator equal to one if the group represents students identified as at risk of dropping out prior to high school, and let Post_t be an indicator equal to one if the cohort entered ninth grade in 2005 or later. The interaction $\text{At Risk}_{sr} \times \text{Post}_t$ captures the treatment effect for at-risk students in the post-policy period. The regression model is:

$$\bar{y}_{srt} = \beta_0 \,\text{Post}_t + \beta_1 \,\text{At Risk}_{sr} + \beta_2 (\text{At Risk}_{sr} \times \text{Post}_t) + \bar{\varepsilon}_{srt},$$

where $\bar{y}_{srt}$ denotes the average outcome for each school-by-risk-status group (e.g., course enrollment, graduation, or college attendance). The coefficient of interest, β_2, measures the average treatment effect on the treated (ATT)—that is, the differential change in outcomes for at-risk students following the policy change, relative to non-at-risk students.

This approach assumes that, absent the policy, both groups would have followed parallel trends in the outcomes of interest. While this assumption is standard in policy evaluations, it may be violated in the present setting. Differences in pre-policy trajectories, for instance due to changes in student composition or district assignment policies, could lead to deviations from parallel trends. Importantly, evidence of parallel trends in the pre-reform period is not sufficient for identification since parallel trends must also hold in the post-reform period when the treatment occurs. Nonetheless, if at-risk and non–at-risk students followed different trends before the reform, it would raise concerns that the estimated effects from the TWFE model capture pre-existing differences rather than the impact of the policy.

4.2 Robustness: Synthetic Two-Way Fixed Effects

To address potential concerns about the parallel trends assumption in the standard TWFE model, I estimate a synthetic two-way fixed effects (STWFE) model, following the approach of Arkhangelsky et al. (2021). The synthetic TWFE framework reweights both units and time periods so that the pre-treatment trajectories of treated and comparison groups are closely aligned. By doing so, this approach attempts to mitigate potential bias from imperfect parallel pre-trends and offers a more credible estimate of the policy's causal effects.

Two sets of weights are used in this procedure. The unit weights $\hat{\omega}_{sr}$ define a synthetic control group by ensuring that control units match the pre-treatment trends of treated units:

$$\sum_{s=1}^{S} \hat{\omega}_{sr} \,\bar{y}_{s1t} \approx \frac{1}{S} \sum_{s=1}^{S} \bar{y}_{s0t}, \quad \forall t = 1, \ldots, T_{\text{pre}}. \tag{1}$$

The time weights $\hat{\lambda}_t$ construct a synthetic pre-treatment period by balancing the influence of pre- and post-treatment periods using comparison group data:

$$\sum_{t=1}^{T_{\text{pre}}} \hat{\lambda}_t \,\bar{y}_{s0t} \approx \frac{1}{T - T_{\text{pre}}} \sum_{t=T_{\text{pre}}+1}^{T} \bar{y}_{s0t}, \quad \forall s = 1, \ldots, S. \tag{2}$$

Given these weights, the synthetic TWFE estimator solves the following weighted least squares problem:

$$\left(\hat{\tau}^{STWFE}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \sum_{s,r,t} \left(\bar{y}_{srt} - \mu - \alpha_r - \beta_t - \tau \text{At Risk}_{sr} \times \text{Post}_t\right)^2 \hat{\omega}_{sr} \hat{\lambda}_t \right\}. \tag{3}$$

Here, At $\text{Risk}_{sr} \times \text{Post}_t$ is an indicator equal to one if the group is classified as At-Risk and enters high school in or after 2005. The coefficient τ captures the average treatment effect on the treated (ATT) for outcome $\bar{y}_{srt}$ after reweighting both units and time periods.

When all weights are set to one, the estimator reduces to the standard TWFE model. If all time weights are set to one and the at-risk fixed effects are omitted, the $\hat{\omega}_{sr}$ weights correspond to standard synthetic control weights. The key advantage of the synthetic TWFE approach is that it relaxes the need for strict parallel trends by requiring only that there exist weights satisfying equations (1) and (2).

In addition to the baseline STWFE specification, I conduct robustness checks that examine the dynamics of the estimated treatment effects. Specifically, I generate panel event-study figures that trace the evolution of outcomes for at-risk and non–at-risk groups before and after the 2005 policy change. These figures assess the plausibility of the parallel trends assumption by evaluating how well the synthetic weights align pre-treatment trajectories. Together, these analyses strengthen the credibility of the main results by showing that they are not driven by apparent pre-trends. Details of this approach can be found in Appendix A.3.

4.3 Results

Table 3 presents results from the standard two-way fixed effects model using pre-high school At-Risk status as the treatment indicator. The policy led to a 1.7 percentage point decrease in low-track assignment and a 6.5 percentage point increase in high-track assignment, indicating a clear shift toward more rigorous curricula for At-Risk students. Dropout rates declined by 4.7 percentage points, and Algebra 2 enrollment increased by 3.6 percentage points. Achievement scores fell by 0.124 standard deviations, implying that the combination of (1) more demanding coursework and (2) a change in peer effects leads to declines in standardized test scores. Standard TWFE estimates imply a 2.6 percentage point increase in overall postsecondary attendance, but this result is driven by a rise in 2-year college enrollment, as 4-year attendance rates fall by 1.6 percentage points for these students.

Table 3: TWFE Results

	(1) Low Track	(2) High Track	(3) Dropout	(4) Achievement Score	(5) Algebra 2 Enrollment	(6) Postsecondary Attendance	(7) 4-Year Attendance
Treatment × Post	-0.0173	0.0645	-0.0473	-0.1240	0.0359	0.0256	-0.0163
	(0.0015)	(0.0018)	(0.0013)	(0.0046)	(0.0017)	(0.0023)	(0.0019)
Observations	801,502	801,502	801,502	697,234	551,380	801,502	801,502
Clustering	District	District	District	District	District	District	District

Standard errors in parentheses.

Table 4 reports the synthetic TWFE estimates, which reweigh treated and control groups to match pre-treatment trends. The results are qualitatively consistent with the standard TWFE estimates in Table 3, though the magnitudes differ for a few outcomes. The estimated decline in low-track assignment is larger at 3.4 percentage points, while the increase in high-track assignment rises to 8.6 percentage points. The estimated reduction in dropout rates is smaller, at 3.5 percentage points, though still statistically different than zero. Algebra 2 enrollment increases by 5.3 percentage points, and the estimated decline in achievement is nearly identical to the standard TWFE estimate, at roughly 0.12 standard deviations.

Table 4: Synthetic TWFE Results

	(1) Low Track	(2) High Track	(3) Dropout	(4) Achievement Score	(5) Algebra 2 Enrollment	(6) Postsecondary Attendance	(8) 4-Year Attendance
Treatment × Post	-0.0344	0.0859	-0.0349	-0.1233	0.0532	0.0200	-0.0175
	(0.0053)	(0.0060)	(0.0041)	(0.0106)	(0.0066)	(0.0040)	(0.0028)
Observations	5,295	5,295	5,295	5,295	4,935	5,295	5,295
Clustering	District	District	District	District	District	District	District

Standard errors in parentheses.

Consistent with the standard TWFE results, the synthetic TWFE estimates indicate a modest increase in overall postsecondary attendance (2.0 percentage points), alongside a 1.8 percentage point decline in 4-year college attendance. Taken together, the results suggest that the main findings are robust and not driven by differences in pre-policy trends, as the synthetic weighting approach yields similar qualitative patterns but slightly larger effects on course-taking and track completion rates.

4.4 The Need for a Model

My design-based estimates provide credible evidence that the 2005 Texas curriculum reform altered track placement, course-taking patterns, and postsecondary enrollment, but they leave several important questions that are best addressed with a structural model. In particular, these estimates capture the net effect of the policy without disentangling the mechanisms driving the observed changes. In this setting, both schools and students make strategic decisions: schools choose track assignments in response to policy incentives (especially in the pre-period), while students adjust course selection, and postsecondary plans based on their assigned curriculum throughout the entire period. The design-based regressions cannot isolate the relative contribution of each side's behavior to the overall policy effect.

Second, the policy changes the distribution of students across standards, directly altering the composition of peers within classrooms. These shifts can influence student achievement, course-taking, and postsecondary plans through direct peer effects, and may also operate indirectly by shaping teacher expectations and instructional practices. These spillovers can work in opposing directions: reduced teacher expectations may lower achievement for some students, while stronger positive peer influences can raise outcomes for lower-performing students. The design-based estimates capture these forces only as part of the overall effect, without separating their individual contributions.

Third, the design-based approach is tied to the specific institutional change implemented in 2005 and cannot speak directly to how outcomes would change under alternative policy designs, such as further increasing curricular rigor, relaxing requirements, or modifying school's incentives. Just as importantly, the policy is unlikely to have affected all students in the same way. Some students may have benefited from exposure to a more rigorous curriculum, while others may have been harmed through this change in within-school peer groups.

With these issues in mind, it is clear that answering additional questions about this policy change requires constructing a model that can be used to run a range of counterfactual exercises. The next section develops such a model, formalizing the joint decision-making process of schools and students to better understand the mechanisms behind the design-based results and to evaluate the effects of alternative policy designs.

5 Model

This section describes a two-stage school model. The first stage adapts Fu and Mehta (2018) to the high school setting: schools make curriculum standards assignments, and students make a course choice given their assignment. Based on this first stage, a student realizes their achievement score and their high school graduation outcome. Given these realizations, students then enter a second stage, where they choose to attend a 2-year college, 4-year college, or no college. I begin by presenting the preliminaries and timing of the model, and then present the details of the model in reverse order.

5.1 Preliminaries and Timing

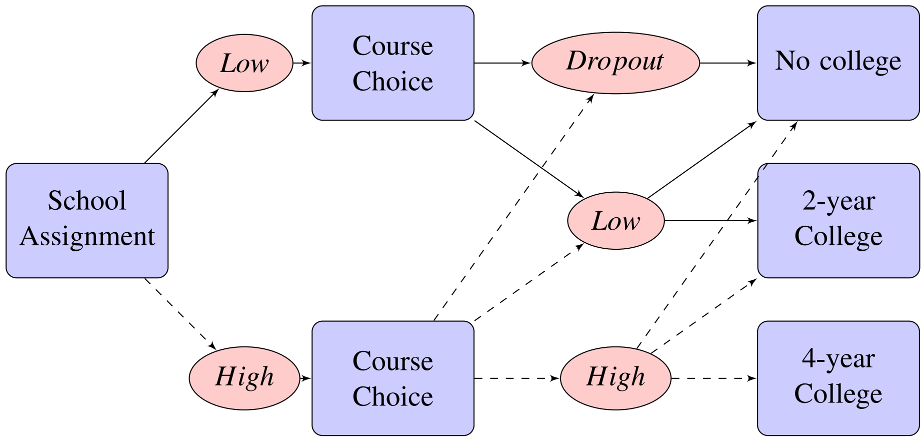


Figure 5: Model Timing

Let $i = 1, \ldots, I$ index students and $s = 1, \ldots, S$ index schools. Each school s is populated by a continuum of students of measure one. Students differ by ability, denoted a_i,[14] and by at-risk status prior to high school, denoted $r_i \in \{0, 1\}$. Both characteristics are observed by the student and the school. Each school s has its own composition of at-risk and non–at-risk students, with the distribution of these subgroups given by $f_{sr}(a)$.

Each school assigns students to one of two curriculum standards, with $j \in \{L, H\}$ denoting the assigned track (low or high). Graduation outcomes are denoted by $g \in \{L, H, D\}$, corresponding to graduation from the low track, graduation from the high track, or dropout.[15] Students make two types of choices: high school course choices (k) and college enrollment decisions (ℓ). The model proceeds in the following stages:

1. Stage 1: Given the mix of students enrolled, schools choose a standards regime, and assign students to curriculum standards. After observing their school's assignment decision, students make a course choice k.

   (a) Students realize their test score y_{ijks} and their graduation outcome g.

2. Stage 2: After realizing their test scores and graduation outcomes, students then make a college choice ℓ.

[14] As it relates to my setting in Texas, ability is measured using students' mathematics standardized test scores from middle school.

[15] This model is focused on the immediate outcomes of high school students. Therefore, a student who eventually receives a GED will be considered a dropout in my model.

Figure 5 maps out the timing of the model graphically. Similar to the instituitional setting I explore in Texas, this model imposes a few restrictions on students pathways. Students assigned to the low standard can either graduate from the low standard or dropout. In addition, those who graduate from the low standard may choose to enroll in a 2-year college, or forego college entirely. High school dropouts are only permitted to choose the no college option.[16]

5.2 College Choice

After making a high school course choice k, the student realizes their standardized test score y_{ijks} and their graduation standard outcome g. With this information, the student then makes a college enrollment decision. If a student graduates from the high standard, they may choose to attend a 2-year college, 4-year college, or no college. If graduating from the low standard, a student cannot attend a 4-year college. The flow utility from selecting college option ℓ is given by:

$$\begin{aligned} U^{uni}_{ijgk\ell s} &= \nu_{jgk\ell} + \lambda_{1g\ell} a_i + \lambda_{2g\ell} y_{ijks} + \lambda_{3g\ell} q_{gs} + \eta^{uni}_{ijgk\ell s} \\ &= u^{uni}_{ijgk\ell s} + \eta^{uni}_{ijgk\ell s}. \end{aligned} \tag{4}$$

In addition to their collegiate options, students have the outside option of not enrolling in college, which yields utility:

$$U^{uni}_{ijgk0s} = \eta^{uni}_{ijgk0s}. \tag{5}$$

The student's level of utility from choosing a given college option ℓ (captured by $\nu_{jgk\ell}$) depends on a student's course choice, their originally assigned standard, and their realized graduation outcome. This is meant to capture the differing levels of disutility students may face upon entering college, given their heterogenous experiences in high school. For instance, depending on their graduating track and performance, students may face more demanding coursework in college or be required to complete remedial classes before advancing to college-level instruction. Moreover, if a student who was initially assigned to the high track but ultimately graduated from the low track (which occurs when $j \neq g$), their previous mix of peers and course requirements may influence their college decision in ways that differ from students who remained in the low track throughout high school.

I also allow preferences to be heterogeneous across student and peer characteristics. Student preferences vary with both their initial ability and their realized achievement, with the latter serving as a signal of their accumulated human capital. Preferences are also shaped by the average ability

[16] In this model, I restrict attention to choices made immediately after high school. In practice, some dropouts may later obtain a GED and subsequently enroll in college, but I shut off this channel.

of a student's peers within their graduated track (q_{gs}), which may influence a student's decision to pursue college, either through encouragement or perceived norms.

5.3 High School Graduation

I assume that students transition from an assigned curriculum standard to a graduation outcome through a Markov process. Specifically, the probability that a student graduates with standard g is given by:

$$P(g \mid j, a_i, r_i, k, q_{js}, y_{ijks}). \tag{6}$$

A student's graduation outcome depends on their initially assigned curriculum track j, their ability a_i, their at-risk status r_i, their course choice k, and the average ability of their peers in the assigned track q_{js}. I estimate these prior to estimating the structural model parameters. A discussion of how I estimate these probabilities can be found in Appendix A.4.

5.4 Achievement Production Function

In addition to observing their high school graduation outcome, a student also learns their standardized test score[17] before making a college enrollment decision. The standardized test score of student i in standard j at school s is governed by the following process:

$$\begin{aligned} y_{ijks} &= \alpha_j + \alpha_{jk} + \alpha_1 a_i + \alpha_2 q_{js} + \alpha_3 \mathbb{1}\{a_i > q_{js}\} + \alpha_4 \mathbb{1}\{a_i > q_{js}\} q_{js} + \varepsilon_{ijks} \\ &= Y(a_i, q_{js}, j, s, k) + \varepsilon_{ijks}. \end{aligned} \tag{7}$$

Achievement scores are shaped by both school assignment decisions and student choices. Schools influence achievement directly through curriculum assignments and indirectly through peer composition. The average track-level ability of peers (q_{js}) captures the overall difficulty level of a student's coursework.

A student's own ability a_i contributes directly to their achievement. A student's own ability fully interacted with peer quality allows for a particular learning environment to affect stronger and weaker students differently. Additionally, a student's course choice k, which is partly determined by the school's curriculum assignment, also affects their preparedness for the achievement test.

[17] Similar to ability, achievement is measured using students' final mathematics standardized test score, taken in 11th grade as part of the Texas Assessment of Knowledge and Skills (TAKS).

5.5 High School Course Choice

Upon realizing their standards assignment, a student is faced with making high school course choices. These choices represent investments in human capital that directly affect test performance. Students assigned to the low standard may either fulfill the minimum requirements of that track or exceed them. Students assigned to the high standard likewise may fulfill or exceed their requirements, but they also have the option to defy their requirements, which results in either graduating under the low standard or dropping out. The flow utility for student i assigned to standard j at school s to invest in course bundle k is:

$$\begin{aligned} U^{hs}_{ijks} &= \kappa_{jk} + \delta_{1jk} a_i + \delta_{2jk} E[y_{ijks}] + \eta^{hs}_{ijks} \\ &\equiv u^{hs}_{ijks} + \eta^{hs}_{ijks}. \end{aligned} \tag{8}$$

A student can also choose their outside option which yields utility:

$$U^{hs}_{ij0s} = \eta^{hs}_{ij0s}. \tag{9}$$

A student's baseline level of utility from selecting k (captured by the parameter κ_{jk}) depends on their initially assigned standard j. This captures heterogeneity in preferences across students in different tracks, reflecting the influence of their curriculum assignments.

In addition, students' preferences over course bundles k may vary with their ability and forecasted achievement ($E[y_{ijks}]$). Students are assumed to have full knowledge of the achievement production function and can anticipate the effect of any investment decision k on their test performance, up to the idiosyncratic test error term ε_{ijks}. This modeling choice allows students to weigh the disutility associated a given course bundle (captured by κ_{jk}) against the potential achievement gains they may receive.

5.6 High School Curriculum Assignment

Prior to the start of high school, each school s assigns its incoming students to one of two curriculum standards: low (L) or high (H). Assignment decisions depend on a student's at-risk status, denoted by $r \in \{0,1\}$. Let $\mu_{jrb} \in [0,1]$ represent the fraction of students with at-risk status r in bin $b \in \{0 \text{ (Below Average)}, 1 \text{ (Above Average)}\}$ who are assigned to standard $j \in \{L,H\}$. By construction, for each combination of r and b, it must hold that $\mu_{Lrb} + \mu_{Hrb} = 1$.[18]

A school's assignment policy can therefore be summarized by the vector $\mu = [\mu_{L00}, \mu_{L01}, \mu_{L10}, \mu_{L11}]$, which specifies the share of non–at-risk and at-risk students assigned to the low standard. For con-

[18] I do not allow schools to initially assign students to dropout.

venience, let $m \equiv m(\mu)$ denote the index corresponding to a particular assignment policy. Given a policy μ, the value to school s is

$$V_s(\mu) = \left\{ \sum_{r,j} \int \left(\omega_1 E[y_{ijks}] + \omega_2 E[c_{ijgk\ell s}] \right) \mu_{jrb(i)} f_{sr}(i)\, di + \phi' \mu + \eta_{sm(\mu)} \right\} \equiv v_s(\mu) + \eta_{sm(\mu)}, \tag{10}$$

$$n_{js} = \sum_r \int \mu_{jrb(i)} f_{sr}(i)\, di,$$

$$q_{js} = \frac{1}{n_{js}} \sum_r \int \mu_{jrb(i)}\, i\, f_{sr}(i)\, di.$$

A school's preferences over a given policy μ depend on how it influences the cohort's expected achievement, $E[y_{ijks}]$, and postsecondary enrollment, $E[c_{ijgk\ell s}]$. These trade-offs are governed by ω_1 and ω_2, which represent the school's relative valuation of achievement and college-going outcomes, shaped in part by the state's accountability system. The parameter vector ϕ captures the school's mean utility from implementing a particular assignment policy.

Under policy μ, the school realizes the share of students in each track, n_{js}, as well as the average peer ability within a track, q_{js}. Any change in μ alters these peer groups, which in turn affects both achievement and college-going outcomes. As a result, schools face a trade-off when choosing μ: they can emphasize higher test performance, greater postsecondary attendance, or attempt to balance the two objectives, depending on the relative magnitudes of ω_1 and ω_2.

5.7 Model Solution

5.7.1 Student Solution

In order to solve both the student and school problem, I assume that all preference shocks are distributed Type-I Extreme Value. I solve the model using backward induction, starting with the student's terminal postsecondary enrollment decision. In this final stage, the student chooses a college option by solving:

$$\max_{\ell} \left\{ u^{uni}_{ijgk\ell s} + \eta^{uni}_{ijgk\ell s} \right\}. \tag{11}$$

In a student's first stage, the value of any given course choice is defined by the following:

$$v^{hs}_{ijks} = u^{hs}_{ijks} + \beta\gamma + \beta \sum_g P(g \mid j, a_i, r_i, k, q_{js}, y_{ijks}) \text{Emax}(a_i, q_{js}, y_{ijks}, j, k, g) \tag{12}$$

where β is the discount factor,[19] and γ is Euler's constant, and $\text{Emax}(a_i, q_{gs}, j, k, g)$ is the expected maximum utility from the college enrollment stage, conditional on graduating under standard g. Since students are not fully aware of their achievement score until the college enrollment stage, the continuation value includes an outer expectation over the test score shock, ε_{ijks}:

$$\text{Emax}(a_i, q_{gs}, y_{ijks}, j, k, g) = \text{E}_{\varepsilon_{ijks}}\left[\log\left(\sum_{\ell}\exp\left(u^{uni}_{ijgk\ell s} + \eta^{uni}_{ijgk\ell s}\right)\right)\right] \tag{13}$$

Given this value function, a student then decides which of the course options is optimal by solving:

$$\max_{k}\left\{v^{hs}_{ijks} + \eta^{hs}_{ijks}\right\}. \tag{14}$$

5.7.2 School Solution

I assume that schools are fully aware of student's preferences and transition probabilities. In addition, I assume that school preference shocks $\eta_{m(\mu)}$ are Type-I Extreme Value and i.i.d. across schools. A school optimally selects a standards regime μ by solving the following problem:

$$\max_{\mu}\left\{v_s(\mu) + \eta_{sm(\mu)}\right\} \tag{15}$$

6 Estimation

6.1 Parameters estimated outside of the model

Before estimating the structural parameters governing student and school preferences, I first estimate the parameters of the achievement production function, $\{\alpha_j, \alpha_k, \alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, outside the main model. These parameters are obtained by estimating Equation (7) using the sample constructed in the TSP data.

6.2 Structural Estimation Overview

6.2.1 Two-Step Estimation

Given the model's timing, and the conditional independence between student and school decisions, I can estimate this model in two steps. In the first step, I estimate all student-side preference parameters, including $\{\kappa_{jk}, \delta_{1jk}, \delta_{2jk}, \nu_{jgk\ell}, \lambda_{1g\ell}, \lambda_{2g\ell}, \lambda_{3g\ell}\}$, using data on course choices and college

[19] I set the discount factor to be 0.95.

enrollment outcomes. In the second step, I estimate the school preference parameters $\{\omega_1, \omega_2, \phi\}$, treating the student-side parameters as given.

6.2.2 Indirect Inference

For both of the stages mentioned in Section 6.2.1, the parameters from this model are estimated by Indirect Inference, a generalized form of the Method of Simulated Moments. Indirect inference minimizes the distance between two sets of auxiliary models estimates, one set from the simulated data that I generate in the economic model, and another from the real data coming from the TSP. I use Indirect Inference rather than maximum likelihood estimation for two main reasons. First, the computational and software demands of maximum likelihood estimation exceed the capacities available on TSP servers. Second, Indirect Inference allows me to link the structural model directly to the design-based estimates presented in Section 4, ensuring that the model parameters are identified by the same variation generated by the policy.

The objective function I minimize with indirect inference is the following:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \left[\hat{\Lambda} - \hat{\Lambda}(\Theta)\right]' W \left[\hat{\Lambda} - \hat{\Lambda}(\Theta)\right] \tag{16}$$

$$\hat{\Lambda}(\Theta) = \frac{1}{H}\sum_{h=1}^{H} \hat{\Lambda}_h(\Theta)$$

where W is the weighting matrix, $\hat{\Lambda}$ is the set of auxiliary parameters estimates coming from the real data, and $\hat{\Lambda}(\Theta)$ is the set of auxiliary parameters estimates coming from my H simulated data sets. $\hat{\Lambda}(\Theta)$ is estimated by first computing the auxiliary parameters for each of the H simulated data sets, then by second taking the average across all H simulated data sets.[20] The asymptotic distribution of the Indirect Inference estimator is:

$$\sqrt{N}(\hat{\Theta} - \Theta_0) \sim N\left(0, (1+1/H)(Q'WQ)^{-1}(Q'W\Omega WQ)(Q'WQ)^{-1}\right).$$

Here, $Q = \frac{\partial \hat{\Lambda}(\Theta)}{\partial \Theta'}$, and Ω is the optimal weighting matrix.[21] When I estimate this model, I set $W = \Omega^{-1}$.

In practice, estimating a dynamic discrete choice model via Indirect Inference is difficult, as discrete choices often result in a non-smooth objective function. To solve this issue, I estimate this model by indirect inference using importance sampling weights, a method suggested by Sauer and

[20] When testing my estimation code, I have found that having a single data set with a sample size approximately equal to the real data set's sample size to be sufficient for recovering the true structural parameters.

[21] Using 1,000 bootstrap samples, I estimate the covariance matrix of auxiliary parameters, and then get an estimate of Ω by taking the inverse of the covariance matrix.

Taber (2021). The procedure for estimation is as follows:

1. With an initial guess of structural parameters Θ_0, solve the model, and simulate a data set from the economic model.

2. Evaluate each individual's conditional choice probability given Θ_0 with the simulated data set. Call it ℓ_{i0}.

3. With a new guess of structural parameters Θ_1, solve the model, but do not simulate. Evaluate each individual's conditional choice probability given Θ_1 with the original simulated data set. Call it ℓ_{i1}.

4. Estimate the auxiliary parameters of this model solved by structural parameters Θ_1, using the ratio ℓ_{i1}/ℓ_{i0} for each individual as an importance sampling weight.

5. Repeat Steps 3 - 4 until the distance between the auxiliary parameters for the real and simulated data are minimized.

6. Repeat Step 1 with the structural parameters that solved Step 5. From there, repeat Steps 2 - 5. Continue this process as needed.

This use of importance sampling weights smooths the objective function, allowing me to apply gradient-based optimization algorithms, such as L-BFGS.

6.3 Identification and Auxiliary Models

6.3.1 Student-side Parameters

Table 5 presents a summary of the moments and auxiliary models I use to identify the structural parameters. In the estimation of student-side parameters, I utilize moments from TSP to identify preferences underlying students' decisions regarding high school coursework and their subsequent college enrollment. These parameters capture (i) mean utilities associated with curriculum and postsondary enrollment decisions, (ii) how individual abilities shape educational decisions, (iii) the influence of academic achievement, and (iv) the role of peer ability in educational outcomes.

Identification of mean utility parameters (κ_{jk} and $\nu_{jgk\ell}$) are identified by the shares $\mathrm{E}[\mathbb{1}_k \mid j]$ and $\mathrm{E}[\mathbb{1}_\ell \mid g,k]$. In the TSP, these reflect the share of each high school course choice (college enrollment choice) conditional on the assigned (graduating) track. Each curriculum standard and course investment choice leads to distinct educational trajectories. Therefore, variation in these conditional probabilities directly reflects differences in baseline utilities associated with each choice.

The effect of initial ability (δ_{1jk} and $\lambda_{1g\ell}$) is identified using observed mean abilities conditional on curriculum and track choices, $\mathrm{E}[a_i \mid k,j]$ and $\mathrm{E}[a_i \mid \ell,g]$. The variation in mean student ability

Table 5: Summary of Identification

Parameters		Targeted Moments
Student-side	Mean Utilities (κ_{jk}, $\nu_{jgk\ell}$)	Share of choices, cond. on assigned/graduating standard
	Initial Ability (δ_{1jk}, $\lambda_{1g\ell}$)	Average ability, cond. on standard and choice
	Academic Achievement (δ_{2jk}, $\lambda_{2g\ell}$)	Average achievement, cond. on standard and choice
	Peer Ability ($\lambda_{3g\ell}$)	Average peer ability, cond. on track and choice
	Additional moments capturing dynamics	Share of college choices, cond. on assigned standard
School-side	Mean Utilities of Policy (ϕ_1, ϕ_2, ϕ_3, ϕ_4)	TWFE regressions from design-based analysis Outcomes: low and high track graduation rate, postsecondary attendance
	Academic Achievement and College Attendance (ω_1, ω_2)	Regression of college attendance on achievement and low track share

into different curriculum and postsecondary choices reflect ability-driven preferences of various education choices.

The role of academic achievement (δ_{2jk} and $\lambda_{2g\ell}$) is identified through observed average achievement outcomes, $\mathrm{E}[y_{ijks} \mid k, j]$ and $\mathrm{E}[y_{ijks} \mid \ell, g]$. Variation in achievement across curriculum and enrollment decisions directly captures how expected performance shapes student preferences.

The effect of peer quality on college enrollment decisions ($\lambda_{3g\ell}$) is identified through observed variation in a student's average graduating track peer ability, $\mathbb{E}[q_{js} \mid \ell, g]$. Differences in peer ability across schools provide the variation needed to identify how a student's high school peers influence their college enrollment choices.

Finally, the dynamic nature of students' decisions is identified using observed college enrollment outcomes conditional on students' initial curriculum assignments, $\mathrm{E}[\mathbb{1}_\ell \mid j]$.

6.3.2 School-side Parameters

To identify the school-side parameters, I exploit variation generated by the track assignment policy that took effect with the 2005 cohort of ninth graders. For a given set of parameters, I solve the school model, simulate student assignments and outcomes, and then estimate auxiliary regressions designed to replicate the design-based analysis in Section 4:

$$y_{ist} = \beta_0 \text{Post}_t + \beta_1 \text{At Risk}_i + \beta_2 \Big(\text{At Risk}_i \times \text{Post}_t\Big) + \varepsilon_{ist}. \tag{17}$$

I estimate this specification for three outcomes: graduation under the low track, graduation under the high track, and college attendance. While all TWFE coefficients are estimated, I only require the model to match β_2, the average treatment effect on the treated.

To identify a school's preference for achievement relative to their other objectives, I estimate the following auxiliary model of school-level college attendance on average achievement and the low-track graduation share:

$$c_s = \beta_0 + \beta_1 y_s + \beta_2 n_{Ls} + \xi_s. \tag{18}$$

This auxiliary model helps quantify the trade-offs schools make across competing goals, providing additional identifying variation for the school's objective function.

6.4 Estimation Results

6.4.1 Achievement Parameter Estimates

Table 6: Achievement Production Function Estimates

	Achievement
Ability	0.546
	(0.00156)
Initial High	-0.234
	(0.00842)
Initial Low $\times$ Exceeds Standard	0.112
	(0.00797)
Initial High $\times$ Completes Standard	0.397
	(0.00552)
Initial High $\times$ Exceeds Standard	0.819
	(0.00571)
Initial Peer Ability	0.218
	(0.00353)
$\mathbb{1}\{$Ability $\geq$ Initial Peer Ability$\}$	0.135
	(0.00274)
Initial Peer Ability $\times$ $\mathbb{1}\{$Ability $\geq$ Initial Peer Ability$\}$	0.135
	(0.00439)
Constant	-0.440
	(0.00660)
Observations	492,425
R-squared	0.6240
σ_ε	0.6130

Standard errors in parentheses.

Table 6 presents estimates related to student performance on their final achievement test. A student's ability at the start of high school is a strong predictor of achievement—an increase of one standard deviation in ability is associated with approximately a 0.5 standard deviation increase in achievement, all else equal.

When students are surrounded by higher-ability peers, schools may respond by increasing course difficulty, which can, in turn, enhance academic performance. My estimates imply that a 1 standard deviation increase in a student's average peer ability increases their achievement by 0.2 standard deviations. I allow this effect to be heterogeneous by student ability and find that students who are more able than their peers benefit even more from an increase in peer quality.

Exceeding the requirements of the low and high curriculum standards is associated with substantial gains in measured achievement. Specifically, exceeding the coursework for the low and high standards corresponds to increases of 0.1 and 0.8 standard deviations in achievement, respectively.

6.4.2 Student-side Structural Parameter Estimates

Table 7 reports the estimated high school course choice parameters for students initially assigned to either the low or high curriculum standard. Students initially assigned to the low standard who choose to exceed their requirements receive higher baseline utility relative to completing their minimum requirements. However, these students are more likely to choose to exceed their curriculum standards if they are of higher ability. Students initially assigned to the low track are less likely to exceed their requirements when they anticipate measurable gains in expected achievement.

Columns 2 and 3 of Table 7 present the estimated parameters governing high-track students' utility from their course investment decisions, relative to the outside option of defecting from their assigned standard. Students who choose to complete the high-track requirements receive a large positive baseline utility, and their utility increases modestly with ability. Expected achievement also strongly contributes to utility, suggesting that students who anticipate academic success find completing the high standard particularly rewarding. In contrast, students who choose to exceed the high-track requirements receive a much lower baseline utility, but the returns to both ability and expected achievement are large.

Table 8 reports the estimated college enrollment parameters for students who graduated with the low or high curriculum standard.[22] For students graduating under the low curriculum standard, the mean utility estimates indicate that all college options are unattractive relative to the outside option of not attending college, suggesting that low track graduates have limited access in continuing into postsecondary education. A student's own ability has a modest positive effect on 2-year college enrollment, while their academic achievement has a negative effect. In addition, having a higher mix of track-level peers discourages these students from attending a 2-year college.

[22] Students who drop out are unable to even attend a 2-year college in my model, so they have no choice to make in the second stage.

Table 7: Parameter Estimates for Students (Course Choice)

	Initial Low Track	Initial High Track	
	Exceed Requirements	Complete Requirements	Exceed Requirements
Mean Utility (κ_{jk})	0.0264 (0.2032)	2.0275 (0.2099)	0.3582 (0.1451)
Ability (δ_{1jk})	0.1906 (0.0786)	0.1547 (0.0554)	0.4821 (0.2365)
Expected Achievement (δ_{2jk})	-0.1251 (0.0688)	0.6601 (0.3535)	2.0874 (0.862)

Table 8: Parameter Estimates for Students (College Choice)

<table>
<tr><th></th><th colspan="2">Low Track Graduate</th><th colspan="2">High Track Graduate</th></tr>
<tr><th></th><th colspan="2">2-Year</th><th>2-Year</th><th>4-Year</th></tr>
<tr><th></th><th>Low Track Initial</th><th>High Track Initial</th><th></th><th></th></tr>
<tr><td>Mean Utility: Enroll in Minimum Coursework ($\nu_{jgk\ell}$)</td><td>-0.5746
(0.1606)</td><td>-1.0135
(0.9453)</td><td>-3.1558
(0.7532)</td><td></td></tr>
<tr><td>Mean Utility: Enroll in College Pre-Req's ($\nu_{jgk\ell}$)</td><td>-0.5689
(0.1495)</td><td>-0.0012
(0.4159)</td><td>-0.5289
(0.0404)</td><td>-2.5018
(0.04500)</td></tr>
<tr><td>Mean Utility: Exceed College Pre-Req's ($\nu_{jgk\ell}$)</td><td></td><td>–2.2421
(1.8827)</td><td>0.2959
(0.0606)</td><td>0.1760
(0.0372)</td></tr>
<tr><td>Ability ($\lambda_{1g\ell}$)</td><td colspan="2">0.3989
(0.0713)</td><td>-0.1782
(0.0796)</td><td>-1.1895
(0.1204)</td></tr>
<tr><td>Achievement ($\lambda_{2g\ell}$)</td><td colspan="2">-0.2925
(0.0145)</td><td>-0.0786
(0.0182)</td><td>3.2564
(0.0686)</td></tr>
<tr><td>Peers ($\lambda_{3g\ell}$)</td><td colspan="2">-0.1484
(0.0185)</td><td>0.2588
(0.0162)</td><td>4.6578
(0.1161)</td></tr>
</table>

Table 9: Parameter Estimates for Schools

	Estimates
Expected Achievement (ω_1)	12.008 (1.4006)
Expected College Attendance (ω_2)	9.2025 (4.0258)
Low Track Assignment: Non-At-Risk, Below Average (ϕ_1)	-14.0898 (0.8459)
Low Track Assignment: Non-At-Risk, Above Average (ϕ_2)	-17.8152 (0.8670)
Low Track Assignment: At-Risk, Below Average (ϕ_3)	2.6130 (0.5468)
Low Track Assignment: At-Risk, Above Average (ϕ_4)	1.5599 (0.3460)

Columns 2 and 3 of Table 8 present the estimated parameters for students who graduate under the high curriculum standard. The mean utility estimates show that exceeding college-level prerequisites increases the value of postsecondary enrollment, particularly for students enrolling in 4-year institutions. Achievement also encourages enrollment in 4-year colleges, while discouraging 2-year college enrollment. Peer quality plays a significant role as well, with students from stronger peer groups more likely to pursue 4-year colleges. In contrast, enrollment in 2-year institutions is associated with weaker peer environments and is more common among students with lower achievement within the high track group.

6.4.3 School-side Structural Parameter Estimates

Table 9 presents the estimated preference parameters for schools. The results indicate that schools place considerable weight on students' expected achievement, while expected college attendance is also valued but to a somewhat lesser degree. The estimates further reveal heterogeneity in the utility schools derive from assigning students to the low track, depending on students' at-risk status. Schools exhibit disutility from assigning non–at-risk students to the low track, while they receive positive utility from assigning at-risk students to the low track. These patterns suggest that schools balance their achievement and postsecondary objectives against differing incentives tied to student characteristics, treating at-risk and non–at-risk students differently when making track assignments.

6.4.4 Model Fit

My model fits each moment and auxiliary model quite well. Tables 10, 11, and 12 present estimated and simulated moments for students. Table 13 presents estimated and simulated moments for schools. In what follows, I will focus on the places where my model could improve in fit.

Table 10 compares student high school and postsecondary outcomes. For students initially in the low track, the model closely matches the share completing minimum requirements and the share exceeding requirements. The model underpredicts 2-year college enrollment for students who graduate from the low track. For students initially in the high track, the model overpredicts the share completing minimum requirements and underpredicts the share exceeding requirements. The model performs reasonably well in capturing college enrollment patterns for this group, slightly overpredicting 2-year enrollment but underpredicting 4-year college enrollment.

Table 10: Unconditional Student Moments

	Initial Low Track		Initial High Track	
	Data	Model	Data	Model
Share Completing Reqs.	0.5116	0.5085	0.4891	0.5317
Share Exceeding Reqs.	0.4884	0.4915	0.4477	0.3865
	Low Track Graduates		**High Track Graduates**	
	Data	Model	Data	Model
Share 2-Year	0.2750	0.2512	0.3051	0.3256
Share 4-Year			0.3040	0.2588

Table 11: High School Course Choice Moments

	Initial Low Track				Initial High Track					
	Course Choice				Course Choice					
	Complete Requirements		Exceed Requirements		Defy Requirements		Complete Requirements		Exceed Requirements	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Average Ability	-0.6921	-0.7237	-0.5478	-0.6551	-0.7457	-0.8110	-0.2567	-0.3507	0.6758	0.5963
Average Achievement	-0.9222	-1.2835	-0.7450	-1.1412	-1.1120	-1.1515	-0.3683	-0.4481	0.6562	0.5983

Table 12: College Enrollment Moments

	Low Track Graduate College Choice		High Track Graduate College Choice			
	2-Year		2-Year		4-Year	
	Data	Model	Data	Model	Data	Model
Average Ability	-0.4962	-0.5736	0.0070	-0.0200	0.6750	0.6634
Average Achievement	-0.6478	-0.8635	-0.0887	-0.1034	0.6563	0.6624
Average Peer Ability	-0.5436	-0.5652	0.1778	0.1654	0.2957	0.2892

Table 13: School Auxiliary Models

		Data	Model
Regression of college attendance on initial standard assignment and achievement	Constant	0.3274	0.3211
	Initial Standard	0.2583	0.2319
	Achievement	0.1137	0.1405
TWFE Regressions (ATT Estimates)	Low Track Graduation	-0.0344	-0.0346
	High Track Graduation	0.0859	0.0920
	College Attendance	0.0402	0.0335

Table 11 evaluates model fit for conditional moments related to course choice among low and high track students. Overall, the model replicates the main patterns in the data and captures the magnitudes of average ability and achievement across choices. For low track students, simulated average ability is somewhat higher than in the data for those who complete or exceed requirements. The model also slightly underpredicts achievement for both low track options. For high track students, the model closely matches the data for those who defy requirements, modestly underpredicts ability for those who complete requirements, and slightly underpredicts it for those who exceed requirements. Achievement is well captured for students who defy or complete requirements, while the model slightly underpredicts achievement for those who exceed requirements.

Table 12 presents conditional moments for low and high track graduates. For low track graduates who enroll in 2-year colleges, the model closely matches the data, slightly underpredicting average ability, achievement, and peer ability. For high track graduates, the model replicates patterns across both 2- and 4-year enrollment choices.

Table 13 reports the fit of the school model. The first set of estimates corresponds to the auxiliary model (18), which regresses college attendance on initial standard and achievement. The model tracks the qualitative patterns well: students assigned to the high standard are more likely to attend college, and achievement is positively associated with enrollment. The signs and relative sizes of the simulated regression estimates align closely with the data. The second set of estimates reports ATT parameters from two-way fixed effects regressions of graduation and college outcomes on at-risk status before and after the 2005 reform. The model reproduces the negative effect of the policy on low-track graduation and the positive effects on high-track graduation and college attendance, and the estimated magnitudes are similar between the simulated data and observed data.

Overall, the model demonstrates a strong overall fit for both the student- and school-level moments. It accurately captures patterns in student course-taking and college enrollment, while also replicating school assignment decisions across ability, at-risk status, and policy periods. These results are encouraging and provide confidence in the predictions presented in the subsequent counterfactual policy experiements.

7 Comparative Statics

In Section 4, I presented TWFE estimates measuring the effect of a default high standard on a range of high school and postsecondary outcomes. Although these estimates provide substantial evidence that the policy improved outcomes across several dimensions, it remains unclear how peer effects, school heterogeneity, and student heterogeneity contribute to these results. Ideally, one would disentangle these components to better understand their individual roles.

In this section, I focus on the role of peer effects in shaping these outcomes. To isolate their contribution to the TWFE estimates, I simulate the structural model under an environment without peer effects. Specifically, this corresponds to setting the relevant parameters in the achievement production function (Equation (7)) to zero:

$$\alpha_2 = \alpha_3 = \alpha_4 = 0, \tag{19}$$

and setting to zero the parameters in the flow utility function (Equation (4)) that govern student preferences for peer ability:

$$\lambda_{3g\ell} = 0 \quad \forall g, \ell. \tag{20}$$

After shutting off these components of the model, I simulate student decisions while holding the school-side assignment rules fixed. Using these simulated data, I then re-estimate TWFE regressions to quantify how outcomes change in the absence of peer effects. In Table 14, I focus on how this adjustment influence postsecondary attendance.

Table 14: TWFE Results without Peer Effects

	(1) Postsecondary Attendance	(2) 4-Year Attendance
Treatment $\times$ Post	0.012 (0.003)	-0.025 (0.002)
Observations	800	800

Standard errors in parentheses

Relative to the results from Table 4, I find that shutting off peer effects reduces the ATT estimate for overall postsecondary attendance from 0.020 to 0.012, a reduction of roughly 40%. For four-year college attendance, the ATT estimate changes from –0.018 to –0.025, an increase of approximately 40% in magnitude. These results suggest that peer effects play an important role in amplifying the policy's positive influence on postsecondary participation. When peer effects are removed, at-risk students are less likely to pursue postsecondary education overall, and the distribution of those who do attend shifts away from four-year institutions.

These findings show that peer interactions are a key mechanism through which curriculum policies influence long-run educational choices, and that the economic model developed here is essential for capturing the equilibrium interactions underlying the estimated effects of the default high-standard policy. While this analysis isolates one channel of influence, it holds the broader policy environment fixed. In the next section, I use the full structural model to move beyond this partial equilibrium analysis and conduct counterfactual policy simulations.

8 Counterfactual Experiments

With the estimated model, I conduct two policy-relevant counterfactual scenarios. The first is a uniform curriculum standard, where schools are banned from initially assigning the low standard curriculum. The second counterfactual allows schools to implement both standards, but adjusts their incentives—much like a state accountability policy could—by shifting the relative preference onto achievement and college enrollment.

8.1 Requiring High Standard Assignment

In this counterfactual, I require high track assignment across all schools. In the context of the model, this provides schools with a singular choice:

$$\mu^* = [0,0,0,0]. \tag{21}$$

This counterfactual approximates the effects of the 2005 ban on low-track initial assignment. This allows me to disentangle the cohort effect from the policy effect, something that can not be done with standard design-based approaches.

Tables 15 present the results of a counterfactual ban on the low standard curriculum. This essentially requires all schools to teach college preparatory coursework. The table compares observed (reality) and counterfactual outcomes across the full sample (Columns 1–2), students originally in the low track (Columns 3–4), and students originally in the high track (Columns 5–6).

At the aggregate level, enforcing universal high track assignment increases the share of students graduating under the high standard and modestly raises the share completing college preparatory coursework. Specifically, the high track graduation rate rises by about 12 percentage points, and the share of students exceeding college pre-requisites increases by roughly 1 percentage point. College enrollment remains stable, and high school dropout rates fall by around 2 percentage points under the policy.

Among students originally in the low track, the policy generates substantial improvements in academic attainment. The share of these students graduating under the high standard rises mechanically from zero to nearly three-quarters, showing that many were capable of meeting this benchmark once assigned. The share completing college preparatory requirements increases by about 17 percentage points. Postsecondary enrollment also improves markedly, with a 7 percentage point increase in 2-year college attendance and a 5 percentage point increase in 4-year enrollment. At the same time, the dropout rate for this group falls sharply, from 37 percent to 14 percent.

For students originally in the high track, the policy induces some negative changes in observed outcomes. The high track graduation rate falls by about 6 percentage points, while dropout rises by roughly 2 percentage points. Postsecondary enrollment also declines: 2-year college attendance falls by about 3 percentage points, and 4-year enrollment decreases by about 0.1 percentage points.

Table 16 shows that requiring all students to begin in the high standard narrows gaps in academic attainment between at-risk and non-at-risk students. For at-risk students, high standard graduation rises by about 12 percentage points (from 56 to 68 percent), while dropout falls by 6 percentage points. Course-taking improves, with enrollment in college preparatory requirements increasing by 6 percentage points and the share exceeding requirements rising by 1 percentage point. Postsecondary enrollment for this group changes little overall: 2-year attendance increases by less than 1 percentage point, while 4-year enrollment falls by about 1 percentage point. Non-at-risk students also see improvements, though smaller given their higher baseline: high standard graduation increases by 3 percentage points (from 85 to 88 percent), dropout falls marginally, and course-taking improves by 2 percentage points in both meeting and exceeding college requirements. Postsecondary enrollment shifts slightly, with 2-year attendance falling by 3 percentage points and 4-year attendance rising by 2 percentage points. These results indicate that while both groups benefit from the policy, at-risk students experience larger relative gains in graduation and reduced dropout, leading to a narrowing of disparities in high school achievement and postsecondary preparation.

Overall, the results show that the gains from requiring universal high track assignment are concentrated among students who would have otherwise been placed in the low track. These students experience large improvements in graduation under the high standard, greater completion of college preparatory coursework, and substantial increases in postsecondary enrollment, alongside lower dropout rates. By contrast, students who were already in the high track experience modest declines in graduation and college enrollment, as well as slightly higher dropout rates. Taken together, the policy generates net improvements in the aggregate, but its benefits are unevenly distributed, with the strongest effects felt by those who would have been assigned to the low track.

8.2 Changing the School's Objective Function

8.2.1 Shifting Incentives Toward College Attendance

In this counterfactual experiment, I adjust each school's objective function to place greater emphasis on maximizing collegiate attendance. This stands in contrast to the prevailing accountability regime during this period, which prioritized student performance on exit-level achievement exams. Within the model, this counterfactual is implemented by increasing the preference schools assign to college enrollment outcomes, relative to achievement.

Table 15: Simulated Results from Requiring a High Standard

	Aggregate		Low Track		High Track	
	Baseline	Counterfactual	Baseline	Counterfactual	Baseline	Counterfactual
High Track Assigned	0.8474	1.0000	0.0000	1.0000	1.0000	1.0000
Low Track Graduate	0.1464	0.1086	0.6319	0.1269	0.0589	0.1053
High Track Graduate	0.7550	0.8145	0.0000	0.7365	0.8910	0.8285
Dropout	0.0957	0.0769	0.3681	0.1366	0.0501	0.0662
Enrolling in College Pre-Req's	0.5139	0.5465	0.5085	0.6770	0.5149	0.5230
Exceeding College Pre-Req's	0.3514	0.3634	0.0000	0.1712	0.4147	0.3980
No College	0.5018	0.5079	0.8579	0.7361	0.4377	0.4668
2-Year	0.2988	0.2856	0.1421	0.2115	0.3270	0.2989
4-Year	0.1994	0.2065	0.0000	0.0524	0.2353	0.2343

Table 16: Simulated Results from Requiring a High Standard, by At-Risk status

	At-Risk		Non-At-Risk	
	Baseline	Counterfactual	Baseline	Counterfactual
High Track Assigned	0.7503	1.0000	0.8947	1.0000
Low Track Graduate	0.2212	0.1606	0.1099	0.0832
High Track Graduate	0.5587	0.6790	0.8506	0.8805
Dropout	0.2200	0.1604	0.0395	0.0363
Enrolling in College Pre-Requisites	0.6118	0.6733	0.4662	0.4847
Enrolling in more than the College Pre-Requisites	0.1700	0.1797	0.4398	0.4529
Enrolling in No College	0.6513	0.6569	0.4290	0.4453
Enrolling in 2-Year College	0.2618	0.2689	0.3168	0.2837
Enrolling in 4-Year College	0.0869	0.0742	0.2542	0.2710

Table 17 reports simulated outcomes under different values of the structural parameter ω_2. When ω_2 is near zero, the simulated outcomes closely resemble what occurs in reality. As ω_2 increases to match the magnitude of the structural parameter for achievement (ω_1), postsecondary attendance rises slightly, by less than 1 percentage point. When ω_2 becomes very large, postsecondary enrollment rises more substantially, by about 6 percentage points relative to the baseline. These gains in college attendance come alongside small declines in average achievement.

However, these benefits are accompanied with trade-offs. As schools place higher preference on collegiate outcomes, the share of students initially assigned to the high track declinces modestly, while the dropout rate increases by about 3 percentage points. Low track graduation also increases. These patterns suggest that schools respond to the change in incentives by adjusting track assignments in ways that may benefit some students' college prospects but come at the cost of higher dropout rates.

Table 18 highlights that these effects are not uniform across groups. For non-at-risk students, placing greater value on collegiate outcomes produces steady gains in postsecondary attendance, rising by over 7 percentage points when ω_2 is made very large. At-risk students also see increases in college enrollment, but the gains are smaller, amounting to only about 2.5 percentage points at the highest values of ω_2. At the same time, achievement for at-risk students falls noticeably as schools reallocate assignments, and their dropout rates rise. These patterns suggest that while stronger incentives for college attendance can narrow gaps in postsecondary participation, they may also exacerbate disparities in achievement and persistence between at-risk and non-at-risk students.

Overall, the counterfactual illustrates that incentivizing schools to prioritize college attendance can meaningfully increase postsecondary enrollment, but also carries risks of unintended consequences for persistence in high school. These findings reinforce the importance of carefully designing accountability systems that balance short-term performance with long-term educational goals.

Table 17: Simulated Results from Increasing the Value of Collegiate Attendance

	$\omega_2 = 0.0$	$\omega_2 = 9.2$ (Baseline)	$\omega_2 = 100.0$	$\omega_2 = 1000.0$
Postsecondary Attendance	0.4792	0.4871	0.5301	0.5409
Achievement	-0.0087	0.0000	-0.0452	-0.0732
Initial High Track Assignment	0.8399	0.8504	0.7438	0.7012
Dropout Rate	0.0985	0.0970	0.1214	0.1287
Low Track Graduation	0.1599	0.1502	0.1949	0.2194

Table 18: Simulated Results from Increasing the Value of Collegiate Attendance, by At-Risk Status

	$\omega_2 = 0.0$		$\omega_2 = 9.2$ (Baseline)		$\omega_1 = \omega_2 = 100.0$		$\omega_2 = 1000.0$	
	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk
Postsecondary Attendance	0.5565	0.3205	0.5663	0.3247	0.6228	0.3396	0.6365	0.3447
Achievement	0.2656	-0.5721	0.2764	-0.5675	0.2664	-0.6849	0.2437	-0.7238
Initial High Track Assignment	0.8857	0.7460	0.8954	0.7578	0.8362	0.5540	0.7922	0.5144
Dropout Rate	0.0399	0.2187	0.0382	0.2178	0.0487	0.2705	0.0562	0.2776
Low Track Graduation	0.1203	0.2413	0.1120	0.2285	0.1497	0.2877	0.1806	0.2993

8.2.2 Shifting Incentives Toward High School Achievement

In this policy simulation, schools are incentivized to maximize achievement by placing greater preference on test performance in their objective functions. This mirrors accountability systems that judge schools primarily on standardized exam results, rather than on longer-run outcomes like college enrollment.

The results in Table 19 indicate that this shift generates measurable short-term gains. Average achievement rises, dropout rates decline, and a larger share of students are assigned to the high track. Table 20 further shows that the achievement gap between at-risk and non-at-risk students narrows by roughly 10%, magnifying the role of peer composition in reducing disparities. Together, these responses are consistent with schools prioritizing improvements in test-based indicators of success.

At the same time, long-run outcomes are harmed. Postsecondary attendance declines as schools optimize for test scores rather than supporting transitions into higher education. Table 20 makes clear that this trade-off is particularly sharp for non-at-risk students: while they see larger gains in measured achievement, their likelihood of attending college decreases modestly.

Overall, this analysis implies that accountability systems centered narrowly on test scores may produce short-run academic improvements while simultaneously undermining longer-run educational attainment.

Table 19: Simulated Results from Increasing the Value of Achievement

	$\omega_1 = 0.0$	$\omega_1 = 12.0$ (Baseline)	$\omega_1 = 100.0$	$\omega_1 = 1000.0$
Postsecondary Attendance	0.4914	0.4871	0.4720	0.4566
Achievement	-0.0102	0.0000	0.0620	0.0656
Initial High Track Assignment	0.8663	0.8504	0.9726	0.9165
Dropout Rate	0.0949	0.0970	0.0653	0.0666
Low Track Graduation	0.1357	0.1502	0.1949	0.1384

Table 20: Simulated Results from Increasing the Value of Achievement, by At-Risk Status

	$\omega_1 = 0.0$		$\omega_1 = 12.0$ (Baseline)		$\omega_1 = 100.0$		$\omega_1 = 1000.0$	
	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk	Non-At-Risk	At-Risk
Postsecondary Attendance	0.5802	0.3092	0.5663	0.3247	0.5381	0.3361	0.5248	0.3165
Achievement	0.2656	-0.5554	0.2764	-0.5675	0.3146	-0.4566	0.3132	-0.4428
Initial High Track Assignment	0.9446	0.7056	0.8954	0.7578	0.9732	0.9712	0.9316	0.8854
Dropout Rate	0.0276	0.2333	0.0382	0.2178	0.0204	0.1575	0.0220	0.1584
Low Track Graduation	0.0784	0.2533	0.1120	0.2285	0.0622	0.1483	0.1013	0.2146

9 Conclusion

This paper provides new evidence on how curriculum standards and school incentives jointly shape students' educational trajectories. Using variation from a Texas policy that made the college-preparatory curriculum the default for incoming ninth graders, I show that raising default standards reduced dropout by 3 percentage points and increased college attendance by 2 percentage points among low-performing students. To interpret these results and assess broader policy implications, I develop and estimate a dynamic model of curriculum assignment, course-taking, and college enrollment that links school decisions to student outcomes.

First, this paper contributes to research on curriculum standards and the role of high school rigor in promoting college readiness. Prior studies show that more demanding coursework improves academic and labor market outcomes, yet little is known about how mandated standards influence course-taking and postsecondary access. I extend this literature by demonstrating that setting the college-preparatory track as the default raises postsecondary attendance by roughly 12 percentage points for students who would otherwise have been placed in a lower curriculum track, substantially increasing access for those at the bottom of the ability distribution.

Second, this paper contributes to the literature on student tracking and peer effects. A large body of work examines how grouping students by ability affects achievement, but few studies model the joint decision-making of schools and students. By explicitly incorporating school assignment rules and peer composition within a dynamic framework, I show how track placement and peer environments reinforce differences in academic preparation and postsecondary choices. A one standard deviation increase in peer ability raises individual achievement by about 0.2 standard deviations, underscoring the importance of peer spillovers in shaping outcomes.

Finally, this paper adds to research on school accountability and performance-based incentives. Existing work emphasizes that accountability policies can improve achievement but also generate strategic behavior and trade-offs. Through counterfactual simulations, I quantify these trade-offs: policies that prioritize college enrollment raise postsecondary attendance from 48% to 54% but increase dropout from 9% to 12%. Policies focused on test performance improve scores by 0.06 standard deviations, reduce achievement gaps by 10%, and reduce dropout from 9.5% to 6.5%, but lower college enrollment from 48%to 45%. These findings highlight that aligning accountability metrics with long-term educational goals is essential for sustaining gains in access without worsening short-run outcomes.

Looking ahead, the framework developed here offers a foundation for evaluating future education reforms that integrate high school and postsecondary policy. Further work could examine how alternative accountability systems or curriculum pathways affect equity across schools and student subgroups. More broadly, this analysis underscores that the design of curriculum and accountability policies plays a pivotal role in determining who is prepared for, and who ultimately accesses, higher education.

References

Achieve. Closing the Expectations Gap: 2014 Annual Report on the Alignment of State K-12 Policies and Practice with the Demands of College and Careers. Technical report, Achieve, Inc, January 2015. URL https://eric.ed.gov/?id=ED554563. Publication Title: Achieve, Inc. ERIC Number: ED554563.

Achieve. How the States Got Their Rates. Technical report, Achieve Inc, October 2015. URL https://www.achieve.org/files/Achieve_StateGradRate2014.pdf. Publication Title: Achieve, Inc.

Kate Antonovics, Sandra E. Black, Julie Berry Cullen, and Akiva Yonah Meiselman. Patterns, Determinants, and Consequences of Ability Tracking: Evidence from Texas Public Schools. Technical Report w30370, National Bureau of Economic Research, August 2022. URL https://www.nber.org/papers/w30370.

Laura M. Argys, Daniel I. Rees, and Dominic J. Brewer. Detracking America's schools: Equity at zero cost? *Journal of Policy Analysis and Management*, 15(4):623–645, 1996. ISSN 1520-6688. doi: 10.1002/(SICI)1520-6688(199623)15:4<623::AID-PAM7>3.0.CO;2-J. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291520-6688%28199623%2915%3A4%3C623%3A%3AAID-PAM7%3E3.0.CO%3B2-J. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291520-6688%28199623%2915%3A4%3C623%3A%3AAID-PAM7%3E3.0.CO%3B2-J.

Dmitry Arkhangelsky, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. Synthetic Difference-in-Differences. *American Economic Review*, 111(12):4088–4118, December 2021. ISSN 0002-8282. doi: 10.1257/aer.20190159. URL https://pubs.aeaweb.org/doi/10.1257/aer.20190159. Publisher: American Economic Association.

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: a survey. *Machine Learning*, 109(4):719–760, April 2020. ISSN 1573-0565. doi: 10.1007/s10994-020-05877-5. URL https://link.springer.com/article/10.1007/s10994-020-05877-5. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 4 Publisher: Springer US.

Julian R. Betts and Jamie L. Shkolnik. The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review*, 19(1):1–15, February 2000. ISSN 02727757. doi: 10.1016/S0272-7757(98)00044-2. URL https://linkinghub.elsevier.com/retrieve/pii/S0272775798000442.

Bart Cockx, Matteo Picchio, and Stijn Baert. Modeling the effects of grade retention in high school. *Journal of Applied Econometrics*, 34(3):403–424, 2019. ISSN 1099-1255. doi: 10.1002/jae.2670. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2670. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.2670.

Olivier De Groote. Dynamic Effort Choice in High School: Costs and Benefits of an Academic Track. *Journal of Labor Economics*, page 726702, July 2023. ISSN 0734-306X, 1537-5307. doi: 10.1086/726702. URL https://www.journals.uchicago.edu/doi/10.1086/726702.

David J. Deming and David Figlio. Accountability in US Education: Applying Lessons from K-12 Experience to Higher Education. *Journal of Economic Perspectives*, 30(3):33–56, September 2016. ISSN 0895-3309. doi: 10.1257/jep.30.3.33. URL https://www.aeaweb.org/articles?id=10.1257/jep.30.3.33.

David Figlio. Testing, crime and punishment. *Journal of Public Economics*, 90(4):837–851, May 2006. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2005.01.003. URL https://www.sciencedirect.com/science/article/pii/S0047272705000708.

David Figlio and Lawrence S. Getzler. Accountability, Ability and Disability: Gaming the System? In *Improving School Accountability*, volume 14, pages 35–49. Emerald Group Publishing Limited, August 2006. ISBN 978-0-7623-1351-8. doi: 10.1016/S0278-0984(06)14002-X. URL https://www.emerald.com/insight/content/doi/10.1016/S0278-0984(06)14002-X/full/html. ISSN: 0278-0984.

David Figlio and Susanna Loeb. School Accountability. In *Handbook of the Economics of Education*, volume 3, pages 383–421. Elsevier, 2011. ISBN 978-0-444-53429-3. doi: 10.1016/B978-0-444-53429-3.00008-9. URL https://linkinghub.elsevier.com/retrieve/pii/B9780444534293000089.

David Figlio and Maurice E Lucas. Do high grading standards affect student performance? *Journal of Public Economics*, 88(9):1815–1834, August 2004. ISSN 0047-2727. doi: 10.1016/S0047-2727(03)00039-2. URL https://www.sciencedirect.com/science/article/pii/S0047272703000392.

David Figlio and Marianne Page. School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Equality?, December 2000. URL https://papers.ssrn.com/abstract=254016.

Chao Fu and Nirav Mehta. Ability Tracking, School and Parental Effort, and Student Achievement: A Structural Model and Estimation. *Journal of Labor Economics*, 36(4):923–979, October 2018. ISSN 0734-306X, 1537-5307. doi: 10.1086/697559. URL https://www.journals.uchicago.edu/doi/10.1086/697559.

Adam Gamoran. The Variable Effects of High School Tracking. *American Sociological Review*, 57 (6):812–828, 1992. ISSN 0003-1224. doi: 10.2307/2096125. URL https://www.jstor.org/stable/2096125. Publisher: [American Sociological Association, Sage Publications, Inc.].

Pamela Giustinelli and Nicola Pavoni. The evolution of awareness and belief ambiguity in the process of high school track choice. *Review of Economic Dynamics*, 25:93–120, April 2017. ISSN 1094-2025. doi: 10.1016/j.red.2017.01.002. URL https://www.sciencedirect.com/science/article/pii/S1094202517300030.

Claudia Goldin and Lawrence Katz. The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005. Technical Report w12984, National Bureau of Economic Research, Cambridge, MA, March 2007. URL http://www.nber.org/papers/w12984.pdf.

Eric A. Hanushek and Ludger Woessmann. Chapter 2 - The Economics of International Differences in Educational Achievement. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*, volume 3, pages 89–200. Elsevier, January 2011. doi: 10.1016/B978-0-444-53429-3.00002-8. URL https://www.sciencedirect.com/science/article/pii/B9780444534293000028.

John Eric Humphries, Juanna Schrøter Joensen, and Gregory F. Veramendi. Complementarities in High School and College Investments. *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4560152. URL https://www.ssrn.com/abstract=4560152.

Juanna Schrøter Joensen and Helena Skyt Nielsen. Is there a Causal Effect of High School Math on Labor Market Outcomes? *Journal of Human Resources*, 44(1):171–198, January 2009. ISSN 0022-166X, 1548-8004. doi: 10.3368/jhr.44.1.171. URL https://jhr.uwpress.org/content/44/1/171. Publisher: University of Wisconsin Press Section: Articles.

Daniel M. Koretz and Sheila I. Barron. The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS). Technical report, Tel: 310-451-7002; Fax: 310-451-6915; E-mail: order@rand, 1998. ISBN: 9780833026873 ERIC Number: ED428131.

Hugh Macartney. The Dynamic Effects of Educational Accountability. *Journal of Labor Economics*, January 2016. ISSN 0734-306X. doi: 10.1086/682333. URL https://www.journals.uchicago.edu/doi/10.1086/682333. Publisher: University of Chicago Press-Chicago, IL.

Hugh Macartney, Robert McMillan, and Uros Petronijevic. A Quantitative Framework for Analyzing the Distributional Effects of Incentive Schemes, May 2021. URL https://www.nber.org/papers/w28816.

Alastair Muriel and Jeffrey Smith. On Educational Performance Measures. *Fiscal Studies*, 32 (2):187–206, 2011. ISSN 1475-5890. doi: 10.1111/j.1475-5890.2011.00132.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-5890.2011.00132.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-5890.2011.00132.x.

Heather Rose and Julian R. Betts. The Effect of High School Courses on Earnings. *The Review of Economics and Statistics*, 86(2):497–513, 2004. ISSN 0034-6535. URL https://www.jstor.org/stable/3211643. Publisher: The MIT Press.

Robert M. Sauer and Christopher Taber. Understanding women's wage growth using indirect inference with importance sampling. *Journal of Applied Econometrics*, 36(4):453–473, June 2021. ISSN 0883-7252, 1099-1255. doi: 10.1002/jae.2818. URL https://onlinelibrary.wiley.com/doi/10.1002/jae.2818.

TAC. Subchapter D - Graduation Requirements, Beginning with School Year 2001 - 2002 (Texas Administrative Code), 2000. URL https://lawrina.com/us-law/state/texas/regulations/title-19/part-2/chapter-74/subchapter-d/.

TAC. Subchapter E - Graduation Requirements, Beginning with School Year 2004 - 2005 (Texas Administrative Code), 2003. URL https://lawrina.com/us-law/state/texas/regulations/title-19/part-2/chapter-74/subchapter-e/.

TEA. 2022 Comprehensive Biennial Report on Texas Public Schools (Document No. GE23 601 03). Technical report, 2022. URL https://tea.texas.gov/reports-and-data/school-performance/accountability-research/comp-annual-biennial-2022.pdf.

THECB. Closing the Gaps by 2015, 2000. URL http://www.txhighereddata.org/index.cfm?objectid=C92AE980-3208-11E8-90D50050560100A9.

A Appendix

A.1 Predicting Initial Assignment

For students entering high school prior to the 2005 academic year, initial assignment is only partially observed. The TSP's administrative data only reports a student's graduation outcome–whether they graduated from the high track, graduated from the low track, or dropped out–but not their initial track assignment. For students who graduate from the high track, initial assignment can be directly inferred: they must have begun high school on the high curriculum standard. In contrast, for students who either graduated from the low track or dropped out, the data do not directly reveal whether they were initially assigned to the high or low standard.

Although not directly observed, a low track graduate's or dropout's initial assignment can often be inferred from individual characteristics, standardized test performance, and the sequence and rigor of their course-taking in both middle and high school.[23] And while assignment rules vary across schools, and are further complicated by differences in student ability and the timing of course enrollment, this is a setting well-suited for machine learning methods. In particular, since "positive" cases (high track assignees) are observed but "negative" cases (low track assignees) are not explicitly labeled, I can employ a Positive-Unlabeled (PU) Learning approach (Bekker and Davis, 2020) to recover plausible initial assignments.

A.1.1 Details on the PU Learning Procedure

To recover a student's initial assignment, I use Positive-Unlabeled (PU) Learning using Random Forests. In traditional supervised learning, a machine learning technique attempts to classify data into "positive" and "negative" groups. However, there are cases when a data set only reports positive cases, and the remaining unlabeled data could either be classified as positive or negative.

PU Learning adapts the traditional supervised learning by trying to classify unlabeled data as either positive or negative. To classify the unlabeled data, I do the following:

1. Identify reliable negatives: from the unlabeled data, construct a set of observations that are most likely negative.

2. Train a classifier: use supervised learning technique with the positively labeled examples and reliable negatives

[23] According to conversations with district administrators, a student's assigment to the low track depended on their pre-high school performance. In some cases, a student's initial assignment can be inferred by their enrollment in "remedial" coursework in high school (e.g. enrolling in pre-Algebra in 9th grade).

3. Select classifier, predict classifications: select the best classifier from (2), and then predict initial assignments.

I discuss the details of these steps, and the data used in these machine learning exercises, in the following sections.

A.1.2 Data Construction for this PU Learning Method

To recover initial curriculum assignments, I draw from two groups of students: (1) all pre-2005 low track graduates and dropouts (the "unlabeled" group), and (2) all students who entered high school in 2005 or later (the "positive" group), since these students were subject to the policy requiring default assignment to the high standard in ninth grade. For all students, I compile standardized test scores from grades 3 through 11, standardized by test year. I also include detailed course enrollment and completion records for math, science, English, social studies, foreign language, and career and technical education, disaggregated at the grade-by-semester level for both middle and high school. Additional covariates include time-invariant and pre-high school characteristics (e.g., ethnicity, gender, at-risk status, socioeconomic status, and English language learner status) and time-varying school-level characteristics (e.g., student course-taking patterns by grade-year, grade-level demographics, average teacher experience, teacher educational attainment, and teacher salaries).

Given the size and complexity of the dataset, missing values are inevitable. To ensure my machine learning tools can utilize the full support of the data, I one-hot encode all variables– transforming both categorical and continuous variables into vectors of binary indicators. Prior to encoding, I discretize continuous variables into decile-based bins. Any missing values, whether from categorical or discretized continuous variables, are flagged with a dedicated "missing" category to preserve information and avoid sample loss.

A.1.3 Identifying Reliable Negatives

In my setting, reliable negatives are a subset of pre-2005 low-track graduates and dropouts who, with a high degree of confidence, were initially assigned to the low track. To identify these reliable negatives, I apply the following procedure separately for low-track graduates and dropouts:

1. Using k-means clustering, I partition students into c groups, where $c \in 1,\ldots,C$. Clustering is based on the full set of features described in Section A.1.2.

2. I select the optimal number of clusters c by choosing the specification that maximizes the Calinski-Harabasz pseudo-F index.

3. For each resulting cluster, I collapse the underlying data into summary statistics and compare them to the equivalent statistics for students in post-2005 cohorts (the "positive" group). The cluster whose centroid is farthest in Euclidean distance from the post-2005 group is then designated as the reliable negative sample.

A.1.4 Training and Selecting a Classifier

I then construct a random forest model to predict whether a student was initially assigned to the high or low curriculum standard. The dataset is first partitioned into a training sample and a testing sample. Within the training sample, I use 3-fold cross-validation to tune hyperparameters and mitigate overfitting. Specifically, I optimize over the number of trees, the maximum number of variables considered at each split, and the minimum node size.[24]

1. Define a grid of candidate hyperparameter combinations: (1) number of trees, (2) maximum variables per split, and (3) minimum node size.
2. Split the training sample into three folds.
3. For each combination of hyperparameters, train a random forest on two folds and evaluate prediction accuracy on the third. Repeat this process across all folds.
4. Repeat steps 2–3 for all hyperparameter combinations in the grid.

After evaluating all models, I select the hyperparameters that yield the highest classification accuracy in predicting high versus low track assignment. Using this optimal configuration, I retrain the random forest on the full training sample. Finally, I assess model performance by evaluating predictions on the held-out testing sample.

A.1.5 PU Learning Results

Table 21 reports cross-validated performance metrics for a range of hyperparameter configurations estimated on the training sample using the PU learning procedure. Each row corresponds to a distinct random forest model defined by the number of variables considered at each split, the minimum node size, and the total number of trees. Metrics include Cohen's Kappa, overall accuracy, positivity rate, negativity rate,[25] and RMSE. The results indicate that the optimal model uses 52

---

[24] Across all hyperparameter configurations, tree depth is left unrestricted.

[25] Let T (F) represent "true" ("false"), and let P (N) denote "positive" ("negative"). Then:

- Positivity Rate = $TP/(TP+FP)$
- Negativity Rate = $FP/(TP+FP)$
- Sensitivity Rate = $TP/(TP+FN)$

variables per split, a minimum node size of 1, and 300 trees. This model achieved the highest accuracy (0.9531) along with strong performance across other metrics, including a high Cohen's Kappa (0.8735), positivity rate (0.9534), specificity rate (0.8779), and a low RMSE (0.2212).

I then take this model's hyperparameters and fit a random forest on the test sample. Table 22 presents the results from this random forest. The model exhibits strong out-of-sample performance, achieving an accuracy of 0.9730 and a Cohen's Kappa of 0.9292. It maintains high sensitivity (0.9725) and specificity (0.9743), indicating strong classification performance for both high- and low-track students. Predictive precision is also high, with a positive predictive value of 0.9913 and a negative predictive value of 0.9220. The root mean squared error (RMSE) on the test sample is 0.1644, further suggesting that the model generalizes well beyond the training data. Figure 6 list the 10 most important variables for this random forest, sorted by their Gini index values.

- Specificity Rate = $TN/(TP+FN)$

Table 21: Random Forest Cross Validation Results

	Number of Variables per Split	Minimal Node Size	Number of Trees	Cohen's Kappa	Accuracy Rate	Positivity Rate	Specificity Rate	RMSE
1	13	1	50	0.7069	0.8987	0.9718	0.6796	0.3182
2	13	5	50	0.7024	0.8974	0.972	0.6739	0.3203
3	13	10	50	0.7026	0.8972	0.9706	0.6769	0.3206
4	26	1	50	0.798	0.9272	0.9712	0.7954	0.2697
5	26	5	50	0.7865	0.9236	0.9717	0.7792	0.2765
6	26	10	50	0.7823	0.922	0.9705	0.7765	0.2792
7	52	1	50	0.8607	0.9485	0.9726	0.8763	0.2269
8	52	5	50	0.8545	0.9464	0.9733	0.8657	0.2315
9	52	10	50	0.8569	0.9472	0.9728	0.8704	0.2298
10	13	1	100	0.7077	0.8994	0.9739	0.6759	0.3172
11	13	5	100	0.7041	0.8981	0.9726	0.6744	0.3193
12	13	10	100	0.7009	0.8972	0.9732	0.6692	0.3206
13	26	1	100	0.8183	0.9343	0.9737	0.8159	0.2564
14	26	5	100	0.8067	0.9303	0.973	0.8024	0.264
15	26	10	100	0.8007	0.9283	0.9728	0.7949	0.2677
16	52	1	100	0.8674	0.951	0.9744	0.8808	0.2214
17	52	5	100	0.8661	0.9505	0.9745	0.8787	0.2224
18	52	10	100	0.8541	0.9463	0.9737	0.8642	0.2316
19	13	1	200	0.715	0.9015	0.9736	0.6852	0.3138
20	13	5	200	0.7084	0.8996	0.9738	0.6769	0.3169
21	13	10	200	0.6996	0.8969	0.9738	0.6663	0.3211
22	26	1	200	0.8102	0.9316	0.9737	0.8053	0.2616
23	26	5	200	0.8073	0.9307	0.9742	0.8003	0.2633
24	26	10	200	0.8024	0.9291	0.9742	0.794	0.2663
25	52	1	200	0.8704	0.9521	0.9754	0.8822	0.2188
26	52	5	200	0.8652	0.9503	0.9746	0.8771	0.223
27	52	10	200	0.8576	0.9476	0.9746	0.8666	0.2289
28	13	1	300	0.7112	0.9004	0.974	0.6798	0.3155
29	13	5	300	0.7081	0.8994	0.9732	0.6779	0.3172
30	13	10	300	0.6995	0.8967	0.9731	0.6677	0.3213
31	26	1	300	0.8125	0.9324	0.9744	0.8065	0.2599
32	26	5	300	0.8084	0.931	0.9735	0.8033	0.2628
33	26	10	300	0.7968	0.9271	0.9732	0.7889	0.2699
34	52	1	300	0.8735	0.9531	0.9749	0.8879	0.2165
35	52	5	300	0.868	0.9512	0.9746	0.8809	0.2209
36	52	10	300	0.8626	0.9493	0.9743	0.8743	0.2252
37	13	1	400	0.7085	0.8996	0.974	0.6766	0.3168
38	13	5	400	0.7045	0.8984	0.974	0.6718	0.3187
39	13	10	400	0.7044	0.8984	0.9741	0.6713	0.3187
40	26	1	400	0.813	0.9325	0.9739	0.8084	0.2598
41	26	5	400	0.8099	0.9315	0.9736	0.805	0.2618
42	26	10	400	0.7998	0.9283	0.9743	0.7901	0.2678
43	52	1	400	0.8732	0.9531	0.9752	0.8866	0.2166
44	52	5	400	0.8645	0.95	0.9748	0.8757	0.2236
45	52	10	400	0.859	0.9481	0.9745	0.8687	0.2279
46	13	1	500	0.711	0.9003	0.9737	0.6803	0.3157
47	13	5	500	0.7121	0.9008	0.9743	0.6803	0.315
48	13	10	500	0.7021	0.8977	0.9738	0.6693	0.3199
49	26	1	500	0.8167	0.9338	0.9741	0.8129	0.2573
50	26	5	500	0.8114	0.932	0.9739	0.8062	0.2608
51	26	10	500	0.7991	0.9279	0.9735	0.7911	0.2685
52	52	1	500	0.8725	0.9528	0.9753	0.8856	0.2172
53	52	5	500	0.8721	0.9526	0.9746	0.8867	0.2177
54	52	10	500	0.8605	0.9486	0.9742	0.8718	0.2268

Table 22: Random Forest Test Sample Predictions

Metric	Value
Sensitivity	0.9725
Specificity	0.9743
Positivity Rate	0.9913
Negativity Rate	0.9220
Kappa	0.9292
Accuracy	0.9730
RMSE	0.1644

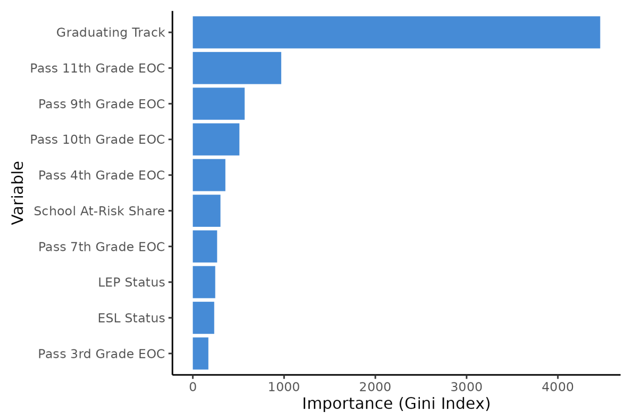


Figure 6: Variable Importance for the Random Forest

A.2 TWFE Results when using Low Tercile as Treatment

Table 23 reports estimates from a standard two-way fixed effects model that compares outcomes for low-performing students before and after the 2005 curriculum policy. The estimates show that making the high-standard curriculum the default led to statistically significant shifts in student outcomes. Assignment to the low track fell by 3 percentage points, while assignment to the high track increased by 4.5 percentage points, suggesting that the policy effectively reallocated students toward more rigorous coursework. Dropout rates declined slightly (-1.6 percentage points), and Algebra 2 enrollment rose by 3.2 percentage points, indicating a modest improvement in course-taking patterns. However, there is a small decline in standardized math scores (-0.032 standard deviations), and postsecondary outcomes appear largely unchanged. Merely looking at postsecondary attendance would suggest that postsecondary outcomes were unaffected, with an estimated effect near zero (-0.0001). However, the policy leads to a modest increase in 2-year college attendance (0.52 percentage points) and a small decline in 4-year college attendance (-0.53 percentage points), suggesting a shift in the type of postsecondary institution attended rather than in overall college enrollment.

Table 23: TWFE Results

	(1) Low Track	(2) High Track	(3) Dropout	(4) Achievement Score	(5) Algebra 2 Enrollment	(6) Postsecondary Attendance	(7) 2-Year Attendance	(8) 4-Year Attendance
Treatment × Post	-0.0295 (0.0015)	0.0452 (0.0018)	-0.0157 (0.0013)	-0.0317 (0.0041)	0.0318 (0.0015)	-0.0001 (0.0024)	0.0052 (0.0023)	-0.0053 (0.0020)
Observations	696,538	696,538	696,538	623,932	486,913	696,538	696,538	696,538
Clustering	District	District	District	District	District	District	District	District

Standard errors in parentheses

Table 24 presents results from the synthetic TWFE estimator, which reweighs groups and time periods to better match pre-treatment trends between treated and control units. Overall, the estimates are consistent in direction and magnitude with the standard TWFE results shown in Table 23, and in most cases, the TWFE point estimates lie within the 95% confidence intervals of the synthetic TWFE estimates. This includes effects on low-track and high-track assignment, Algebra 2 enrollment, and each postsecondary outcome. The only exception is the effect on dropout rates.

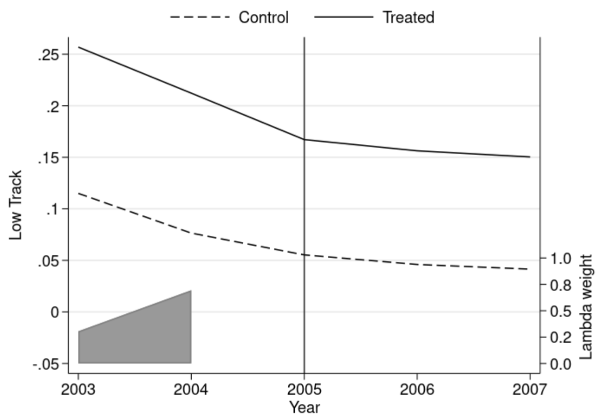
Table 24: Synthetic TWFE Results

	(1) Low Track	(2) High Track	(3) Dropout	(4) Achievement Score	(5) Algebra 2 Enrollment	(6) Postsecondary Attendance	(7) 2-Year Attendance	(8) 4-Year Attendance
Treatment × Post	-0.0258 (0.0052)	0.0471 (0.0060)	0.0038 (0.0049)	-0.0247 (0.0080)	0.0246 (0.0067)	-0.0011 (0.0037)	0.0059 (0.0034)	-0.0064 (0.0025)
Observations	5,005	5,005	5,005	5,005	4,675	5,005	5,005	5,005
Clustering	District	District	District	District	District	District	District	District

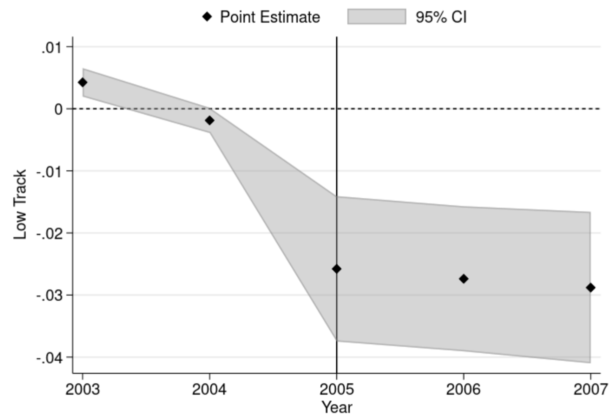
Standard errors in parentheses

A.3 Event Study Results

A.3.1 Treating At-Risk Status as Treated

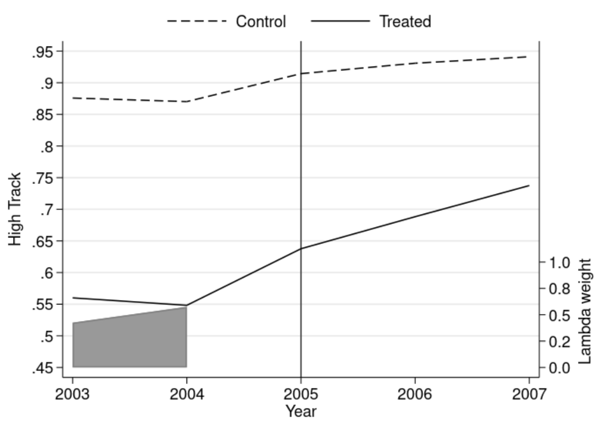


(a) Outcome Trends

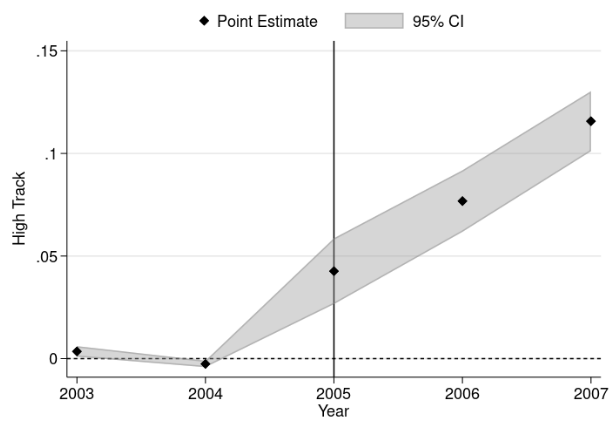


(b) Event Study

Figure 7: Event Study, Low Track

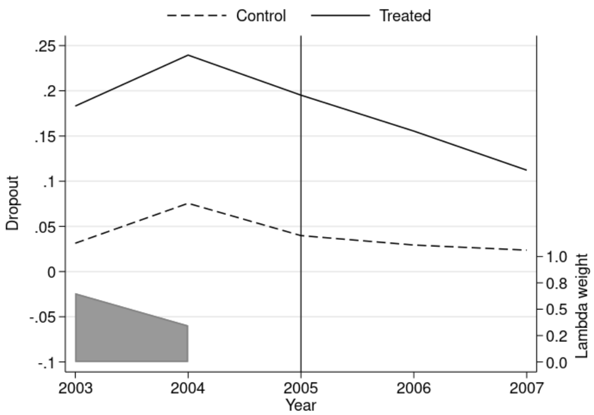


(a) Outcome Trends

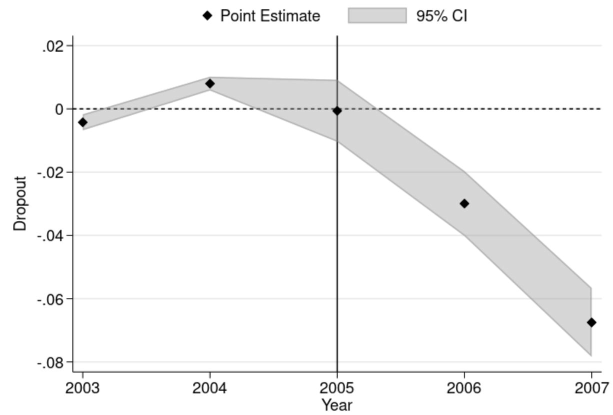


(b) Event Study

Figure 8: Event Study, High Track

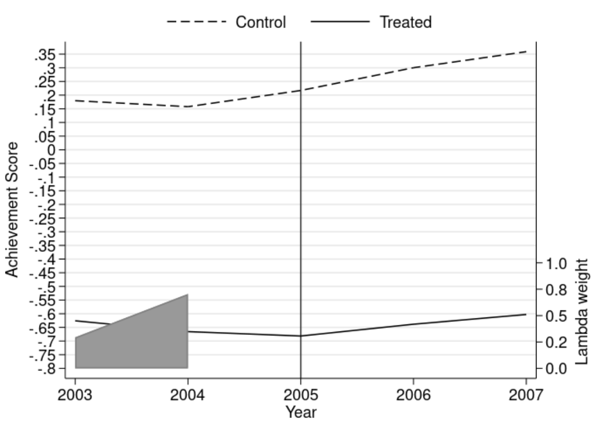


(a) Outcome Trends

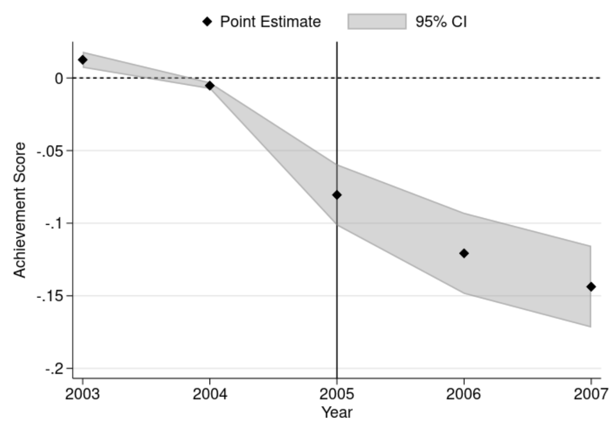


(b) Event Study

Figure 9: Event Study, Dropout

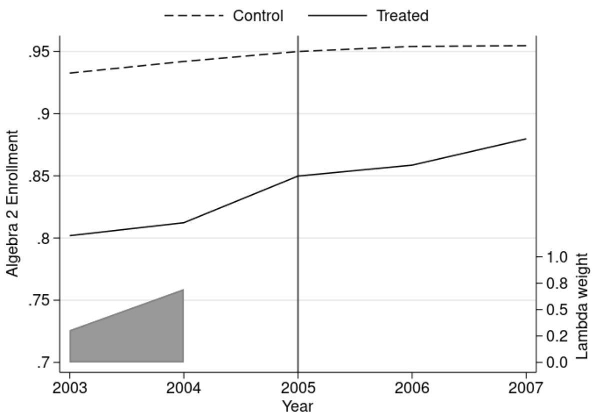


(a) Outcome Trends

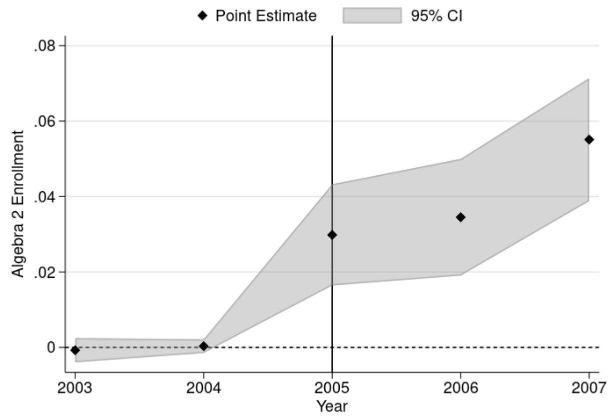


(b) Event Study

Figure 10: Event Study, Eleventh Grade Achievement

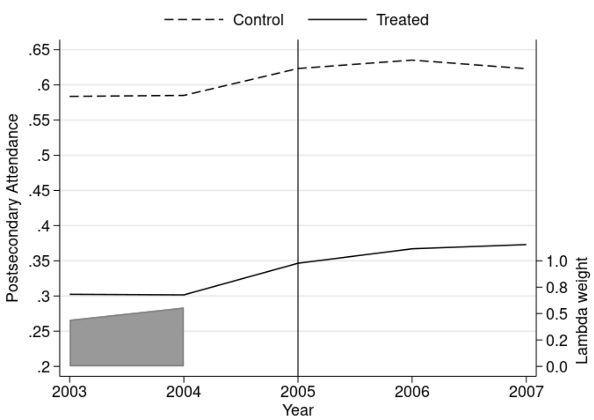


(a) Outcome Trends

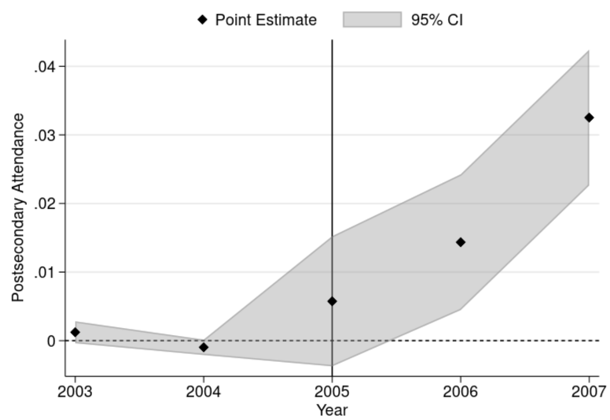


(b) Event Study

Figure 11: Event Study, Algebra 2 Enrollment

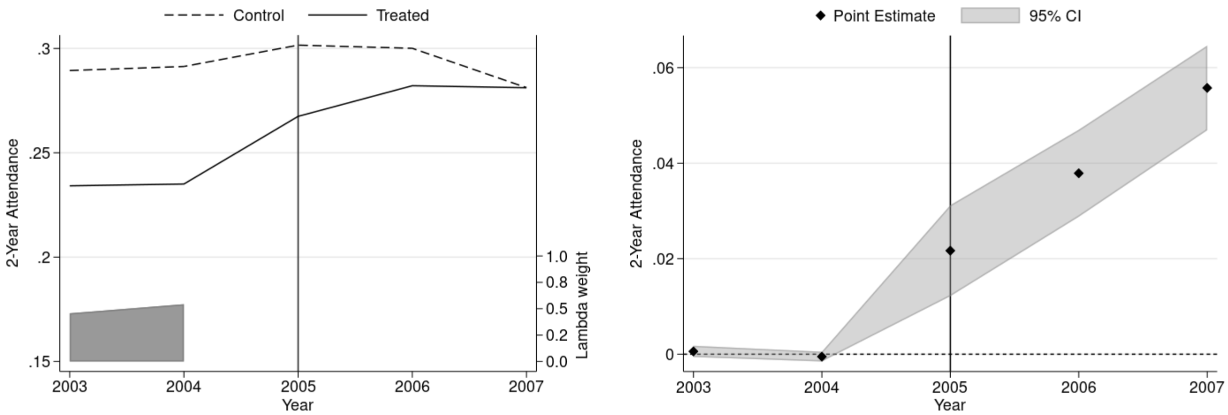


(a) Outcome Trends



(b) Event Study

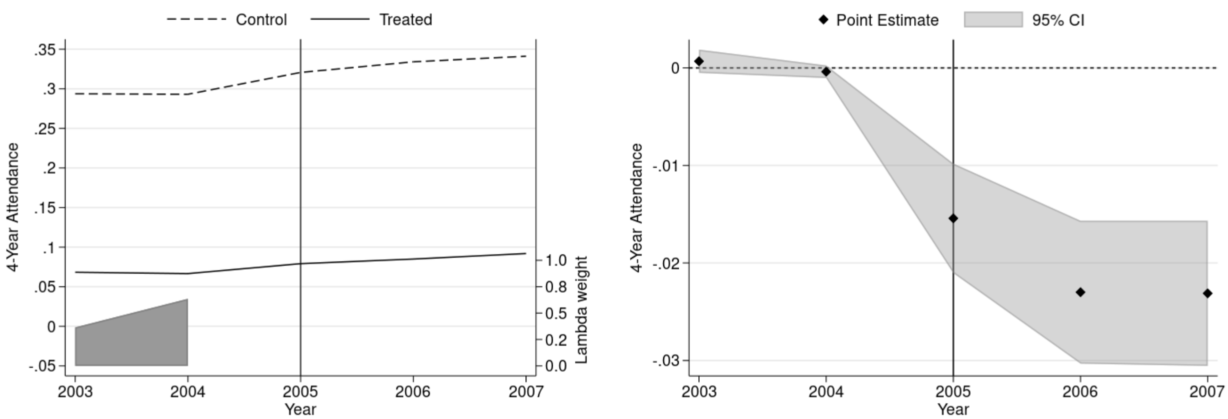
Figure 12: Event Study, Postsecondary Attendance



(a) Outcome Trends

(b) Event Study

Figure 13: Event Study, 2-Year College Attendance

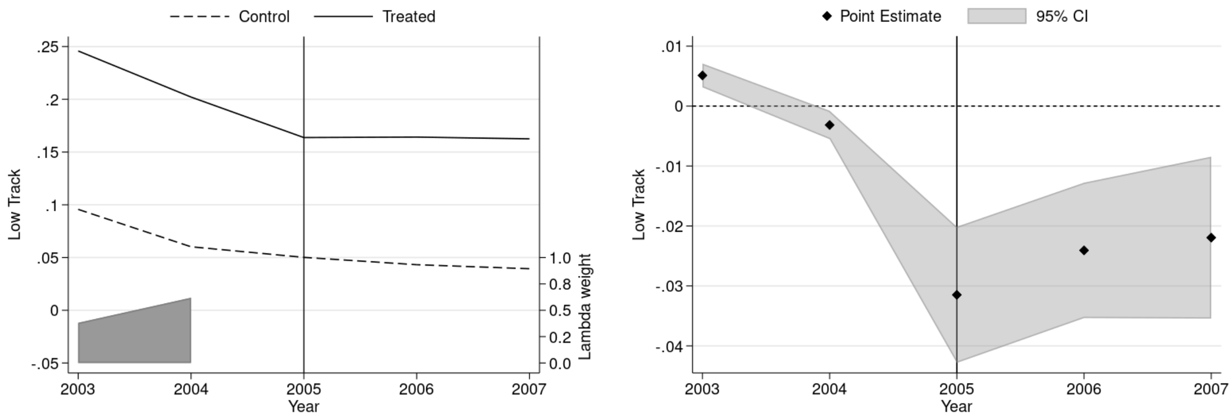


(a) Outcome Trends

(b) Event Study

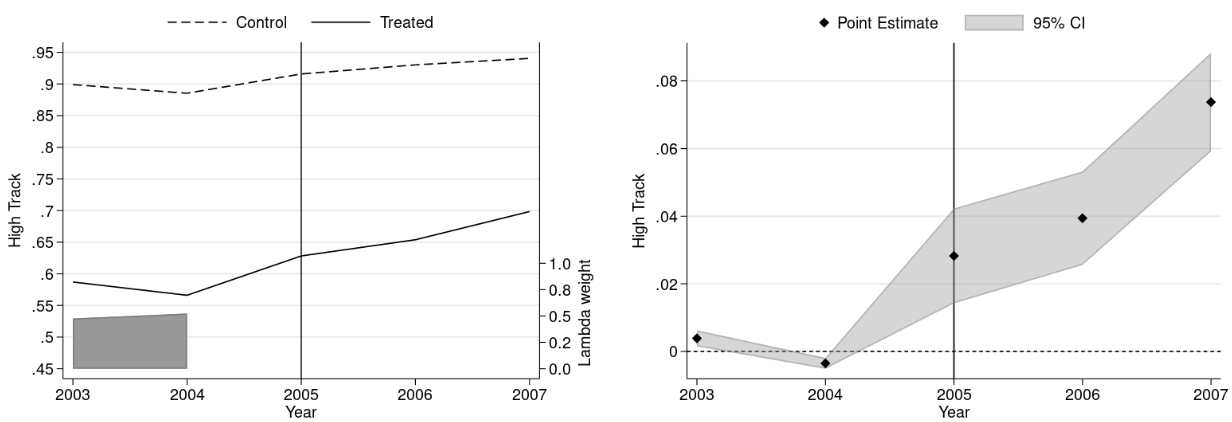
Figure 14: Event Study, 4-Year College Attendance

A.3.2 Treating Low Tercile Students as Treated



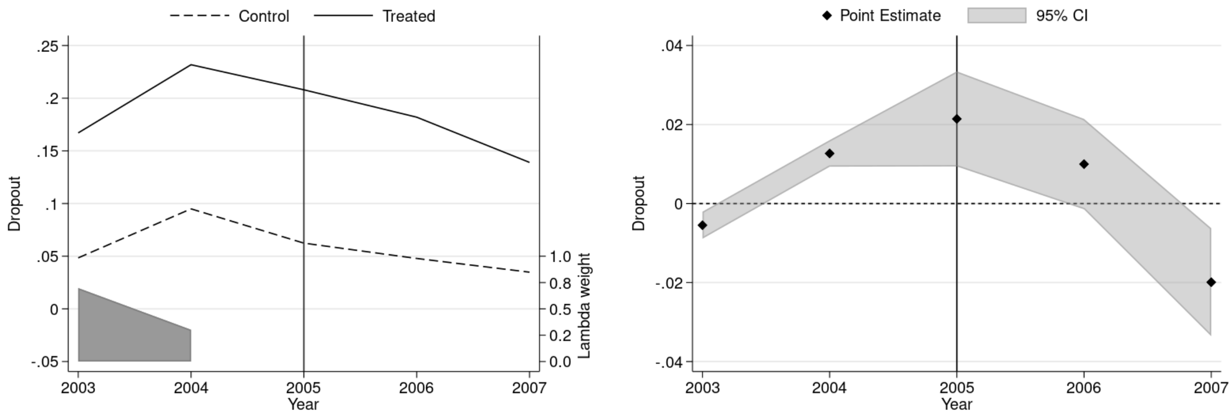
(a) Outcome Trends (b) Event Study

Figure 15: Event Study, Low Track



(a) Outcome Trends (b) Event Study

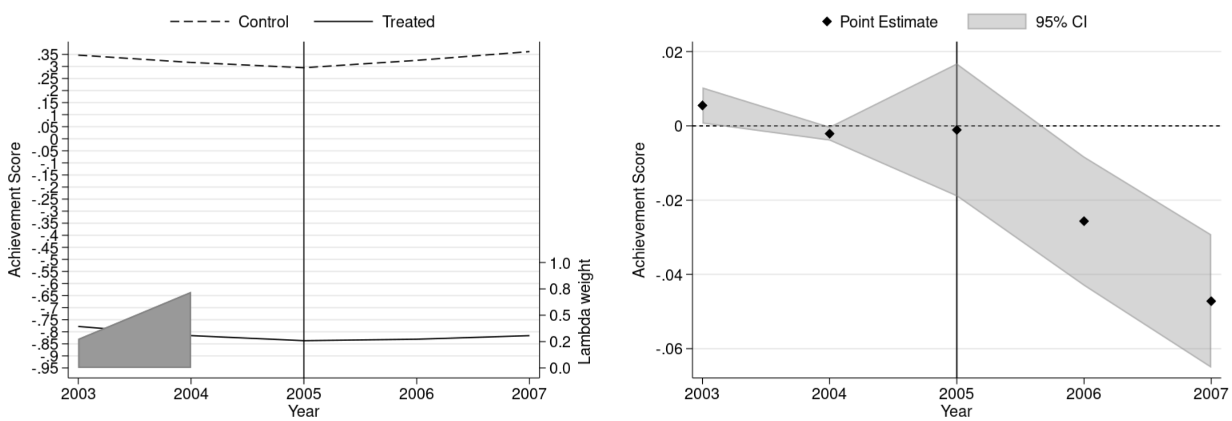
Figure 16: Event Study, High Track



(a) Outcome Trends

(b) Event Study

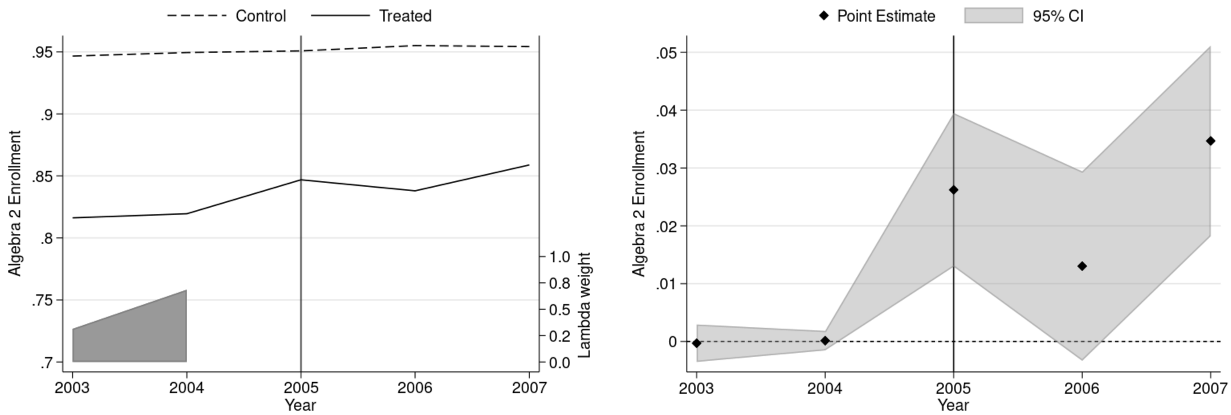
Figure 17: Event Study, Dropout



(a) Outcome Trends

(b) Event Study

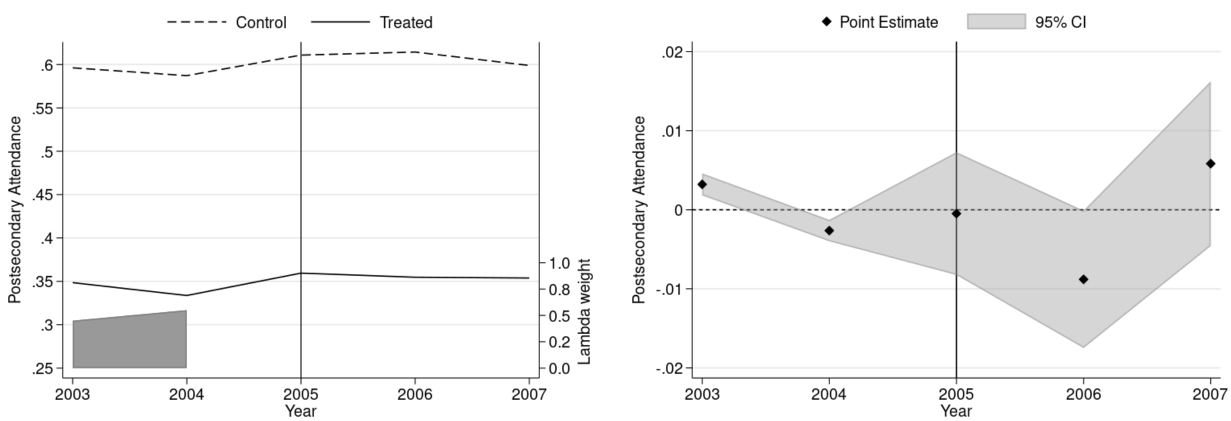
Figure 18: Event Study, Eleventh Grade Achievement



(a) Outcome Trends

(b) Event Study

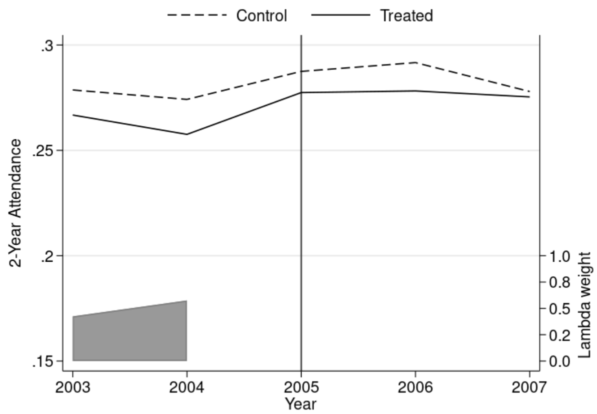
Figure 19: Event Study, Algebra 2 Enrollment



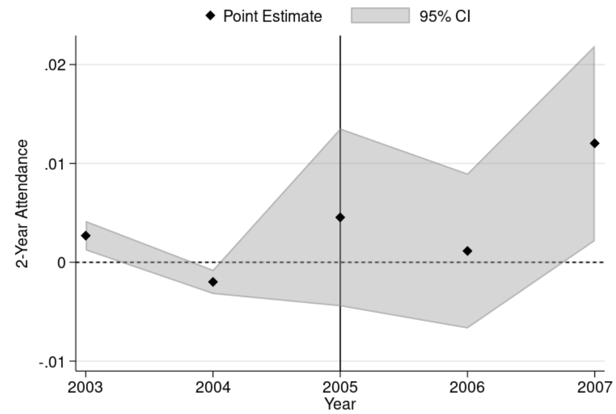
(a) Outcome Trends

(b) Event Study

Figure 20: Event Study, Postsecondary Attendance

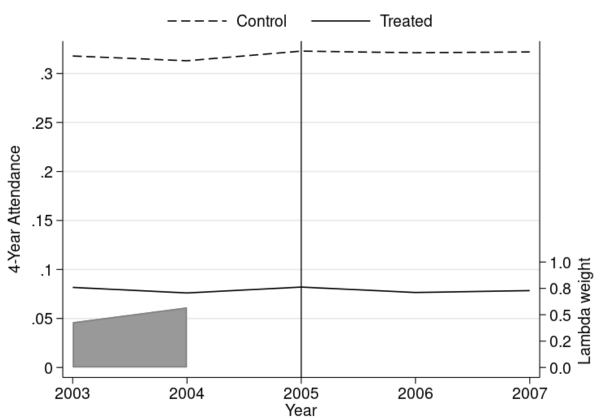


(a) Outcome Trends

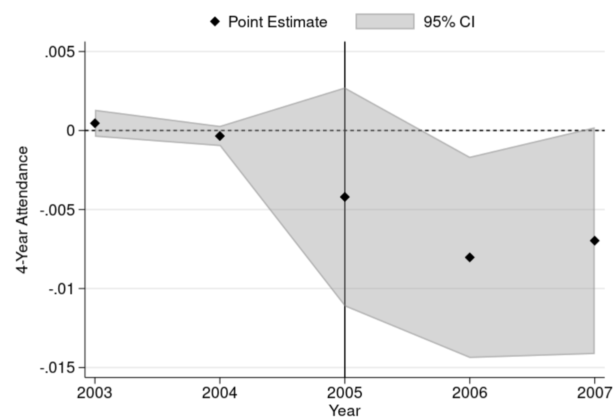


(b) Event Study

Figure 21: Event Study, 2-Year College Attendance



(a) Outcome Trends



(b) Event Study

Figure 22: Event Study, 4-Year College Attendance

A.4 High School Graduation Probabilities

To estimate the transition probabilities discussed in Section 5.3, I adopt a semi-parametric approach. I begin by discretizing each of the six covariates (j, a_i, r_i, k, q_{js}, and y_{ijks}) into bins. Within each resulting cell of the grid, I compute the empirical probability of each possible graduation outcome. In cases where a grid cell contains too few observations to credibly estimate a probability, I instead use fitted values from a fully saturated multinomial logit model.