Ignorance is Bliss? Evaluating the Impact of Rank Non-Disclosure on High School Course Choice*

Ritika Sethi[1]

[1]Department of Economics, Rice University

October 20, 2024

Abstract

Discontinuing the disclosure of class rank to high school students is under consideration at various schools, with arguments suggesting that emphasizing rank may distort the incentives that shape students' course choices. This issue holds particular significance for policymakers in Texas, where class rank directly influences admission to in-state public universities. Since high school courses can have long-term effects on students' educational and labor outcomes, it is crucial to understand the implications of altering students' information during their decision-making process. In this paper, we propose a dynamic model of learning to capture the course choices of high school students. Each year, students receive feedback in the form of interim rank, allowing them to form beliefs about their future rank and subsequently make course choices. Combining results from the estimation of this model and a quasi-experimental analysis, we reveal the significance of rank information in shaping students' decision-making. Additionally, we simulate course choice probabilities under different counterfactual rank disclosure policies.

***Keywords*:**
***JEL Codes*:**

*ritika.sethi@rice.edu, Department of Economics, Rice University, Houston, TX 77005.

1 Introduction

Class rank plays a significant role in college admissions across the United States, particularly in states with percent plan policies that guarantee acceptance to public universities for students graduating above a certain percentile threshold in their high school class. Texas, for example, grants automatic admission to any public university in the state, including prestigious flagship institutions, to students in the top 10% of their class. With such policies directly tying college access to class rank, a student's relative position can have profound long-term impacts on their educational and career trajectories.

A growing literature documents that a higher academic rank in high school positively affects students' subsequent life outcomes (Zax and Rees, 2002; Cicala et al., 2011; Murphy and Weinhardt, 2020; Bertoni and Nisticò, 2023; Denning et al., 2023).[1] Some research also suggests that providing rank information to students can lead to improvements in academic performance, especially for high-achieving students (Azmat and Iriberri, 2010; Goulas and Megalokonomou, 2021; Brade et al., 2022). As a result, understanding how information about rank shapes students' academic choices and preparation is crucial for educators and policymakers.

In recent years, there has been a growing trend among school districts to restrict the disclosure of class rank information to students. Proponents argue that an excessive focus on class rank fosters an unhealthy competitive environment and leads students to make course enrollment decisions aimed at gaming the ranking system rather than pursuing their authentic academic interests. Districts have adopted various policies to limit rank information, such as only reporting rank to the top 10% of students, providing rank in terms of quartile ranges, or fully withholding rank information.

High school course selection can have significant downstream effects on college and career outcomes. Empirical evidence shows that course choices in high school influence college attendance, choice of major, and labor market outcomes (Rose and Betts, 2004; Aughinbaugh, 2012; Arcidiacono et al., 2012; Altonji et al., 2012; Tchuente, 2016; Todd and Wang, 2021). Given the long-term importance of high school course-taking, it is crucial to understand how adjusting the provision of rank information may affect students' course selection behavior.

This study examines the impact of class rank disclosure policies on high school students' course choices, with a focus on mathematics. We develop and estimate a dynamic structural model of course choice under imperfect information, building upon the discrete choice framework of Keane and Wolpin (1997) and Arcidiacono (2004). Our model incorporates Bayesian belief updating about students' own abilities and their cohort's average ability, based on signals from GPA and class rank. Importantly, we allow for potentially biased initial beliefs about academic abilities, which we estimate by leveraging variation in choices among students with similar objective success probabilities but different subjective expectations. The dynamic belief formation process provides a natural framework for evaluating policies that alter the availability of academic performance information.

[1] Delaney and Devereux (2022) offers an excellent review of this literature.

Our empirical analysis uses comprehensive administrative data from a large urban school district in Texas between 2006-2016, covering over 100,000 students across 32 schools. The data includes student demographics, course histories, grades, GPA, and class rank, allowing us to richly model the dynamics of high school course choice.

We estimate the model using the conditional choice probability (CCP) method of Hotz and Miller (1993). This approach leverages the property that some actions are terminal states (e.g. leaving school), simplifying the dynamic problem by expressing continuation values in terms of future choice probabilities. Following James (2011) and Arcidiacono et al. (2025), we integrate out latent abilities and employ an EM algorithm that treats ability as known at the maximization step, yielding a tractable learning model. Building on Arcidiacono and Miller (2011, 2019), we incorporate unobserved heterogeneity, known to the individual, while maintaining computational traceability.

Identification of the model leverages variation in observed choices and outcomes, coupled with assumptions on the heterogeneity and learning processes. We model both permanent unobserved heterogeneity in preferences and abilities as discrete student "types", and continuous course-specific abilities that are initially unknown and learned over time. We attribute "lumpy" variation in choices and outcomes among observationally similar students to type differences, and residual "smooth" variation to ability differences. The learning process is identified from changes in outcomes and choices as students accumulate academic performance signals.

Our estimated model aligns closely with descriptive patterns in the data. Simulations of class rank non-disclosure counterfactuals reveal that the policy effects are larger for higher-ability students in later grades. Removing rank information leads to increases in advanced course-taking among high achievers, consistent with reduced discouragement effects. These findings highlight the role of dynamic considerations and the potential for class rank disclosure to distort the allocation of student effort. To corroborate our structural results, we conduct a quasi-experimental analysis exploiting a natural experiment in class rank disclosure policies across multiple school districts. Using a generalized difference-in-differences design, we find large positive effects of rank non-disclosure on the likelihood of taking advanced math courses, with the effects growing over time. The distributional patterns of the quasi-experimental estimates align with our model's counterfactual predictions, providing an important external validation. Our paper contributes to the literature on educational choice, learning processes, and the impact of relative performance feedback. We innovate by modeling multi-dimensional Bayesian belief updating based on absolute and relative performance signals, bridging the gap between the learning and feedback literatures. Methodologically, our use of unobserved heterogeneity in both flow utility and transitions, estimated via a sequential EM algorithm, pushes the frontier of modeling unobserved heterogeneity in dynamic structural models.

The remainder of the paper proceeds as follows. Section 2 details the related literature. Section 5 outlines our dynamic model of high school course choice, highlighting the role of imperfect information and belief updating. Section 4 describes our data sources. Sections 6 and 7 detail our identification strategy and estimation procedure. Section Section 8 reports the structural parameter estimates and model fit. Section 9 conducts counterfactual simulations of rank non-disclosure policies. Section 10 presents our quasi-experimental analysis and results. Section 11 discusses the implications of our findings and concludes.

2 Related Literature

This section reviews the relevant literature on dynamic discrete choice models in education, learning processes, methodological approaches, and the role of feedback in both educational and contest settings.

Dynamic Discrete Choice in Education: Dynamic discrete choice models have been extensively applied to educational decision-making processes. These models capture the sequential nature of educational choices and account for the uncertainty and learning that occurs over time.

A seminal contribution to this field was made by Keane and Wolpin (1997), who developed and estimated a dynamic structural model of schooling, work, and occupational choice decisions using data from the National Longitudinal Surveys of Labor Market Experience (NLSY). This work laid a strong foundation for subsequent research in the area, demonstrating the potential of dynamic structural models to provide insights into complex life-cycle decisions related to human capital accumulation.

Building on this foundation, Eckstein and Wolpin (1999) examined youth behavior in the context of education and employment choices. Their study focused on the factors influencing high school dropout decisions, incorporating heterogeneity in skills and preferences.

Arcidiacono (2004) made significant contributions by introducing ability sorting into the dynamic framework of college major choices. This work highlighted the importance of considering student heterogeneity and learning about one's abilities in educational decision-making.

More recently, Arcidiacono et al. (2025) extended this line of research by estimating a dynamic structural model where individuals face uncertainty about their academic ability and productivity, which respectively determine their schooling utility and wages. Their framework accounts for heterogeneity in college types and majors, as well as occupational search frictions and work hours. They find that removing informational frictions would increase the college graduation rate.

De Groote (2019) develop a dynamic model in which students choose the academic level of their program and their effort level. They find that encouraging underperforming students to switch to less academic programs reduces grade retention and dropout.

Learning: The process of learning plays a crucial role in both educational and economic decision-making. Several key studies have contributed to our understanding of learning dynamics in various contexts.

Miller (1984) applied learning models to job matching, demonstrating how individuals update their beliefs about job suitability over time. This work has important parallels in educational settings, where students learn about their abilities and preferences for different subjects or majors.

In the context of consumer choice, Erdem and Keane (1996) developed a model of learning under uncertainty. While focused on brand choice, their framework provides valuable insights into how individuals make decisions with imperfect information, a concept highly relevant to educational choices.

Arcidiacono (2004) incorporated learning about abilities into models of educational choice, showing how students update their beliefs about their aptitude for different fields of study based on their academic performance.

More recently, Arcidiacono et al. (2025) extended this line of research by estimating a dynamic structural model in which individuals learn about their academic ability and productivity, which respectively determine their schooling utility and wages.

Methodology: The estimation of dynamic discrete choice models often involves complex computational challenges. Several methodological approaches have been developed to address these issues. Hotz and Miller (1993) introduced the conditional choice probability (CCP) approach, which significantly reduced the computational burden of estimating dynamic discrete choice models. This method has been widely adopted in the education literature. Keane and Wolpin (1997) applied these techniques to career decisions while incorporating a finite mixture model to capture unobserved heterogeneity. Arcidiacono and Miller (2011) further advanced the methodology by developing techniques for estimating dynamic models with unobserved heterogeneity. Their approach combines the CCP method with the expectation-maximization (EM) algorithm, allowing for more flexible modeling of individual differences.

Role of Feedback in Education: Feedback plays a crucial role in shaping educational outcomes and student behavior. Several studies have examined the effects of different types of feedback in educational settings. Azmat and Iriberri (2010) investigated the impact of relative performance feedback on high school students. Their study found that providing students with information about their relative standing improved performance, particularly for those at the extremes of the achievement distribution. Tran and Zeckhauser (2012) examined the effects of rank feedback in a field experiment, finding that knowledge of one's rank can serve as an inherent incentive for improved performance. Floyd et al. (2024) explored the consequences of withholding grades, providing insights into how the absence of explicit feedback affects student behavior and subsequent employment outcomes. Elsner and Isphording (2017) studied the "big fish in a small pond" effect, showing how a student's relative rank within their school influences human capital investment decisions and long-term outcomes. Goulas and Megalokonomou (2021) investigated the short- and longer-term effects of providing students with feedback about their true academic standing. Their findings suggest that accurate self-knowledge can lead to more efficient educational investments.

Role of Feedback in Contests: The literature on feedback in contests provides valuable insights that can be applied to competitive educational settings. Genakos and Pagliero (2012) studied the effects of interim rank feedback in dynamic tournaments. Their findings suggest that relative performance information can influence risk-taking behavior and effort allocation in competitive environments. Lemus and Marshall (2021) examined feedback in dynamic innovation contests, providing insights into how information revelation affects participant strategies and contest outcomes. Jiang et al. (2018) analyzed the role of feedback in crowdsourcing contests, demonstrating how different feedback mechanisms can influence participation and performance in competitive settings.

In conclusion, this paper contributes to and extends several key strands of the above discussed strands of literature. Our work builds upon the foundational models of educational decision-making while focusing on the course choices of high school students. Additionally, we extend these models by incorporating a dual learning process where students update their beliefs about both individual and cohort ability based on GPA and rank signals, respectively, bridging the gap between the literature on learning with studies on the impact of relative performance feedback. Methodologically, our use of unobserved heterogeneity in both flow utility and the transition process, estimated via a sequential EM algorithm, builds on the computational advancements of Arcidiacono and Miller (2011). Finally, our counterfactual analysis examining the impact of removing rank signals on course choices contributes to the ongoing policy debate on the role of comparative feedback in educational settings. By integrating these various elements, our paper provides a comprehensive framework for understanding the complex dynamics of high school course selection, offering insights that can inform both future research and policy discussions in education.

3 Background

Percent Plans In the wake of the abandonment of race-conscious affirmative action policies, percent plans for college admission were adopted in California, Texas and Florida. These plans represent a return to an old method of admitting students to leading colleges - the evaluation of high school grades and class standing.

The Texas Ten Percent Plan, established in 1997, guarantees automatic admission to Texas public universities for students graduating in the top 10% of their high school class. This policy allows eligible students to choose from any of the 35 public universities in Texas, including flagship institutions like the University of Texas at Austin (UT Austin) and Texas A&M University. The high school class rank is used as the primary criterion for admission under this plan. It is calculated by the district or school using the most recent academic record available at the time of application, which could include the student's rank at the end of the 11th grade, mid-12th grade, or at the time of high school graduation.

California's program guarantees admission to one of the University of California (UC) system's eight undergraduate campuses for students who graduate in the top 4% of their high school class. However, this policy does not guarantee admission to the student's campus of choice. Eligibility is based on the student's GPA in specific UC-required courses taken during the 10th and 11th grades. Because ELC students are not guaranteed a spot at a particular UC campus, traditional admission considerations remain in place at each campus. For example, highly selective campuses like UC

Berkeley evaluate applicants using additional criteria such as academic preparation, extracurricular achievements, personal qualities, and other factors.

Florida's program guarantees admission to the state university system for public high school students who graduate in the top 20% of their class, as determined by their district. Students are eligible for admission if they either have a B average in required high school academic units or meet a combination of GPA and admission test scores on a sliding scale. While the program offers entry into the state university system, it does not guarantee placement at a student's preferred university. Eligible students still need to compete for admission to specific institutions, which may have their own additional criteria and performance standards.

In sum, there are several important differences in the percent plans in Texas, California, and Florida. First, which students are eligible in the broadest sense varies and is changing even as the plans are being implemented. Public and private students in Texas and California may be eligible under these states' plans; only public school students in Florida have the same opportunity. Whereas currently, percent-plan-eligible students in Texas have to graduate only with the minimum required credits, soon they will have to meet more strenuous high school coursework requirements to benefit from the 10 percent plan. These plans do not all make similar guarantees to eligible students; California and Florida just promise access to the state university system. Only Texas promises access to the premier institutions, which, similar to Florida, are the only places where raceconscious admission is a salient factor. Finally, the method and data by which eligibility is determined differs. In Texas and Florida, individual districts calculate senior GPA based on all or a particular subset of coursework completed by a designated time. California, however, requires districts, using specified coursework, to identify a larger pool of juniors from which, after parental permission is obtained, the University of California determines the smaller group of senior students eligible in "local context." These and a host of other more subtle differences outlined above represent important caveats. The percent plans in Texas, California, and Florida are not the same in how they are structured and in what they deliver. Any one statement that claims to encompass all percent plans is simply inaccurate.

Before delving into the specifics of the Texas admission plan, it is important to highlight the central role that class rank plays in the Texas education system. Class rank is the sole determinant of eligibility for the Texas Ten Percent Plan. While other supporting materials, such as state testing and SAT scores, are required as part of the overall college application, they do not influence admission under this plan. This unique focus on class rank sets Texas apart from other states and ensures that students' placement within their high school cohort is the primary criterion for guaranteed admission. Another distinguishing feature of the Texas plan is its relatively flexible course and credit requirements, making it more accessible to a broader range of students. Unlike California and Florida, Texas guarantees eligible students admission to its flagship institutions, such as the University of Texas at Austin and Texas A&M University.

Texas Top Ten Percent Plan The Texas Top Ten Percent Plan, also known as Texas House Bill 588, was implemented in 1998 to grant automatic admission to any public university in the state for students who rank within the top ten percent of their high school graduating class. The primary goal of this policy was to increase representation of students from underperforming high schools in selective universities without explicitly relying on race-based affirmative action. Given the economic and racial segregation in Texas public schools, policymakers expected that the plan would enhance diversity by using school-specific class rank as the primary criterion for eligibility.

While House Bill 588 formalized this automatic admission policy, it was not a radical departure from previous practices at the University of Texas (UT). Until 1993, UT allowed automatic admission for students in the top 10 percent of their class. In 1994, the university adjusted its admission criteria to include a combination of class rank and SAT scores, making the process more restrictive. House Bill 588, signed into law by Governor George W. Bush and passed by the 75th Texas Legislature in 1997, reverted to the previous approach by basing eligibility solely on class rank and eliminating the standardized test score requirement for automatic admission. This policy sought to broaden access to highly selective public institutions by enabling students to compete with peers within their own schools, thereby leveling the playing field for students from different educational backgrounds.

Research on the Texas Top Ten Percent Plan has shown varied impacts on college enrollment and long-term outcomes. Using regression discontinuity, Niu and Tienda (2010) finds that the policy increases flagship university enrollment for Hispanic students and those from schools with a high proportion of minority and economically disadvantaged students. Black et al. (2020) further demonstrate that the policy benefits highly ranked students from non-traditional feeder schools, boosting their college enrollment, graduation, and earnings, while students from traditional feeder schools, who lose access to flagship universities, do not experience declines in overall enrollment or long-term earnings. Similarly, Daugherty et al. (2014) show that guaranteed admission significantly increases flagship enrollment and persistence, especially among students in schools with high college-going rates, but has little effect on students from the most disadvantaged schools and results in a displacement from private universities rather than an increase in overall college enrollment.

Some studies suggest that the heterogeneous effects of the Top Ten Percent Plan can be attributed to strategic school choices aimed at maximizing class rank. Cullen et al. (2013) analyze students' transitions between 8th and 10th grades three years before and after the policy change and find that among students with both the motive and opportunity for strategic high school choice, at least 5% enroll in different high schools to enhance their chances of graduating in the top 10%. These students often opt for their neighborhood high schools instead of transferring to more competitive institutions, which, regardless of their own race, can displace minority students from the top 10% pool. Additionally, Cortes et al. (2014) provide evidence that families reacted strategically to this policy by relocating to neighborhoods with lower-performing schools, thereby increasing property values in those areas. This effect is most pronounced in schools that were very low-performing before the policy change. They also find that the strategic reactions were influenced by the availability of local schooling options, with the effects of the Top 10% Plan being weaker in areas with fewer school choices.

Concerns have been raised regarding students manipulating the system through strategic course selection to boost their class rank (The Texas Senate, 2001). Initially, the Texas Top Ten Percent Plan did not specify any required academic curriculum for eligibility, which led to worries that students were opting out of rigorous courses. In response, the Texas legislature passed an amendment in 2001 that increased curriculum requirements for eligibility (Office of House Bill Analysis, 2001). The amendment, effective for the graduating class of 2008 and beyond, introduced three curriculum options: the minimum graduation plan (which excluded college preparatory courses), the state-recommended graduation plan (which included college prep courses), and the advanced graduation plan. Under the new rules, students must complete either the recommended or advanced graduation plans to qualify for automatic admission under the Top Ten Percent Plan.

Additionally, school districts have the authority to determine how class ranks are calculated and disclosed, and in recent years, several districts have revised their rank calculation policies (Webb, 2019b; Dellinger, 2022; Potter, 2018; Donaldson, 2020; Rozen, 2022; Webb, 2019a). These changes often involve adjustments to the eligible courses and the weights assigned to them. Districts also have discretion over how rank information is shared with students. While state law mandates that rank information be disclosed to students within the top ten percent, some districts have proposed alternative methods of disclosure, such as providing rank information in quartiles or withholding specific rank details altogether. One motivation behind these policy changes is the concern that an overemphasis on class rank encourages unhealthy competition among students, leading them to take courses primarily to boost their rank rather than selecting classes that align with their college and career goals. Similar concerns have prompted schools and parents in other states to consider adopting comparable rank disclosure policies (Balingit, 2015; Hui, 2022).

Table 1. Changes in Local Policies

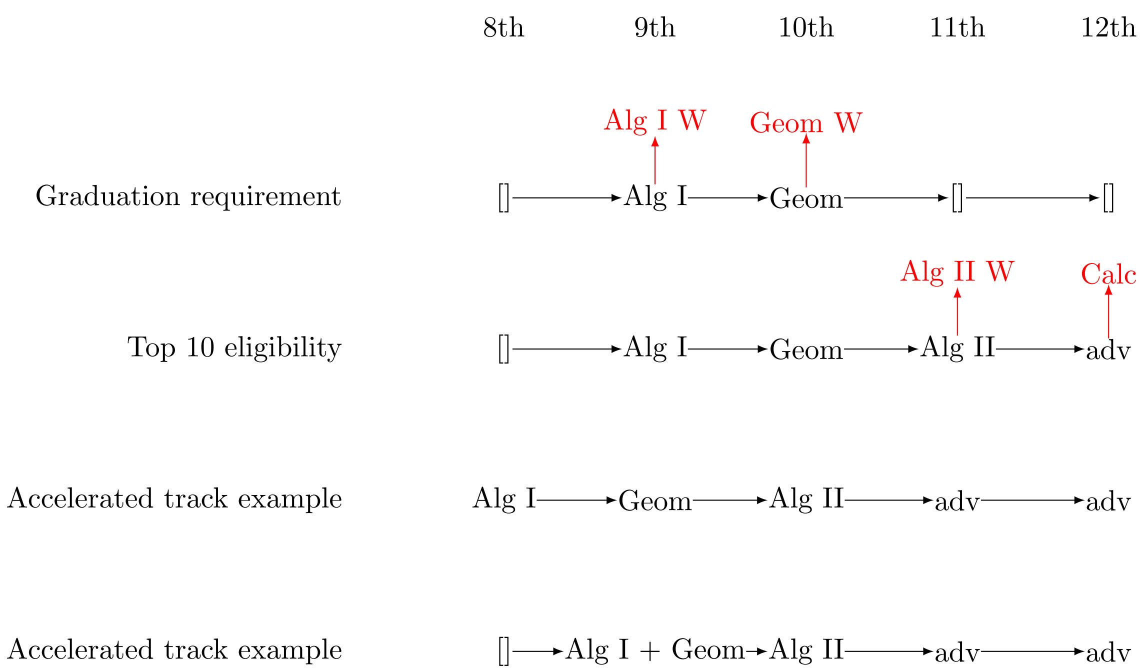
District	Policy	Effective	Change
Dallas	Rank calculation	Class of 2021	Rank will now be based on 15 core courses in English Language Arts, Math, Science, and Social Studies. It was earlier based on any 12 courses including electives.
Frisco	Rank disclosure	Class of 2019	For students outside the top ten percent, rank will not be reported in their transcript. Students will still know their GPA. The lowest GPA in the top 10 percent, first quartile, second quartile and third quartile will be published in the Student Portal.
Clear Creek	Rank calculation	Class of 2027	Rank will be based on only core (non-elective) courses in English Language Arts, Math, Science, Social Studies, and World Languages and Culture.
Katy	Rank disclosure	Class of 2021	For students outside the top ten percent, only quartile rank will be reported in their transcript.

3.1 Math Course-taking in Texas

According to the current high school graduation requirements in Texas, students must complete Algebra I and Geometry. Within these courses, students have the option to choose between regular and weighted versions, such as Pre-AP Algebra I and Pre-AP Geometry. The recommended path is to take Algebra I in the ninth grade and Geometry in the tenth grade. However, students who aspire to be eligible for consideration in the Top Ten Percent Plan need to go a step further and take Algebra II, as well as an advanced math course like Calculus. These advanced courses also offer weighted versions, such as Pre-AP Algebra II and AP Calculus.

Students have additional decisions to make regarding their math course selection. They can opt for the accelerated track, where they take Algebra I in middle school, Geometry in ninth grade, Algebra II in tenth grade, and advanced math courses in the final two years of high school. Another option to expedite their progress is to double up on math courses in the same year, such as taking both Algebra I and Geometry in the ninth grade. Figure 1 provides an overview of the various course sequences available to a student. It is important to note that math courses follow a sequential structure, with Algebra I serving as a prerequisite for Geometry, which is a prerequisite for Algebra II, and so on, culminating in Calculus.

Figure 1. Course map: horizontal and vertical tracking



4 Data

We use administrative data obtained from a large urban school district in Texas. The dataset includes academic records for 106,761 high school students and covers their demographic information (race, ethnicity, gender, and economically disadvantaged status), course selections, course grades, GPA, and class rank. The data spans 32 schools and 304 graduating classes between 2006 and 2016. Additionally, we observe the high school-level courses credited to students during their middle school years (6th, 7th, and 8th grades) as well as during high school (9th through 12th grades).

4.1 Descriptive Statistics

Table 2. Past academic and background characteristics by course choice

	(1) [Min,Max]	(2) Mean	(3) S.D.
Panel A: Demographics			
Black	[0,1]	0.25	(0.44)
Hispanic	[0,1]	0.60	(0.49)
Female	[0,1]	0.50	(0.50)
Panel B: Math Coursetaking			
Total Math courses	[0,5]	3.28	(0.88)
Total Weighted Math courses	[0,4]	0.71	(1.16)
Ever Completed Algebra I	[0,1]	1.00	(0.06)
Ever Completed Geometry	[0,1]	0.93	(0.25)
Ever Completed Algebra II	[0,1]	0.80	(0.40)
Ever Completed Pre-Calculus	[0,1]	0.48	(0.50)
Ever Completed Calculus	[0,1]	0.14	(0.34)
Panel C: Dropout			
Dropout by Grade 9	[0,0]	0.00	(0.00)
Dropout by Grade 10	[0,1]	0.09	(0.28)
Dropout by Grade 11	[0,1]	0.18	(0.38)
Dropout by Grade 12	[0,1]	0.26	(0.44)
Panel D: GPA			
GPA after Grade 9	[70,100]	84.93	(4.73)
GPA after Grade 10	[70,100]	84.95	(4.15)
GPA after Grade 11	[70,100]	85.22	(3.81)
GPA after Grade 12	[70,99]	85.70	(3.58)
Obs		106,761	

Note: This table reports summary statistics for the data that is used to estimate our structural model. This is student-level data.
Data source: Administrative data from a large urban district in Texas

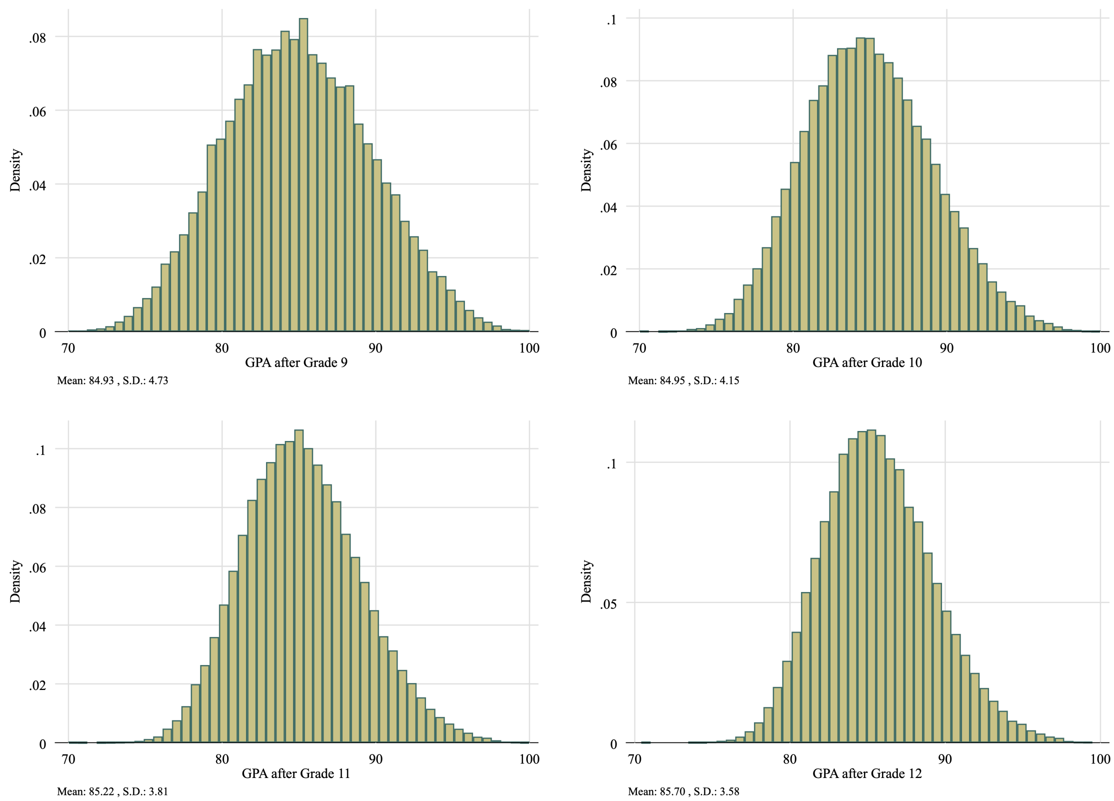
Table 11 provides a summary of the academic and demographic characteristics of the 106,761 students in our dataset. Panel A presents the demographic breakdown of the student sample. On average, 25% of the students identify as Black, while 60% are Hispanic, reflecting the district's diverse student population. Female students make up 50% of the sample.

Panel B summarizes the math course-taking patterns. On average, students completed 3.28 math courses, with 0.71 of those being weighted (e.g., advanced or honors courses). Nearly all students completed Algebra I (100%) and Geometry (93%), while 80% completed Algebra II by the end of high school. However, more advanced courses see a sharp drop in completion rates, with 48% of students taking Pre-Calculus and only 14% completing Calculus.

Panel C reports the dropout rates by grade level. The dropout rate is zero by Grade 9 but increases steadily as students progress through high school. By Grade 10, 9% of students have dropped out, with the rate rising to 18% by Grade 11 and 26% by Grade 12. This indicates a significant attrition rate, particularly in the later years of high school.

Panel D provides GPA statistics over time. Students' average GPA shows a steady increase from 84.93 after Grade 9 to 85.70 by the end of Grade 12. Meanwhile, the standard deviation of GPA decreases from 4.73 in Grade 9 to 3.58 in Grade 12, suggesting reduced variation in academic performance as students approach graduation. Figure 2 further illustrates the distribution of GPA by grade level, showing that GPAs are approximately normally distributed across each grade.

Figure 2. GPA Distribution by Grade Level



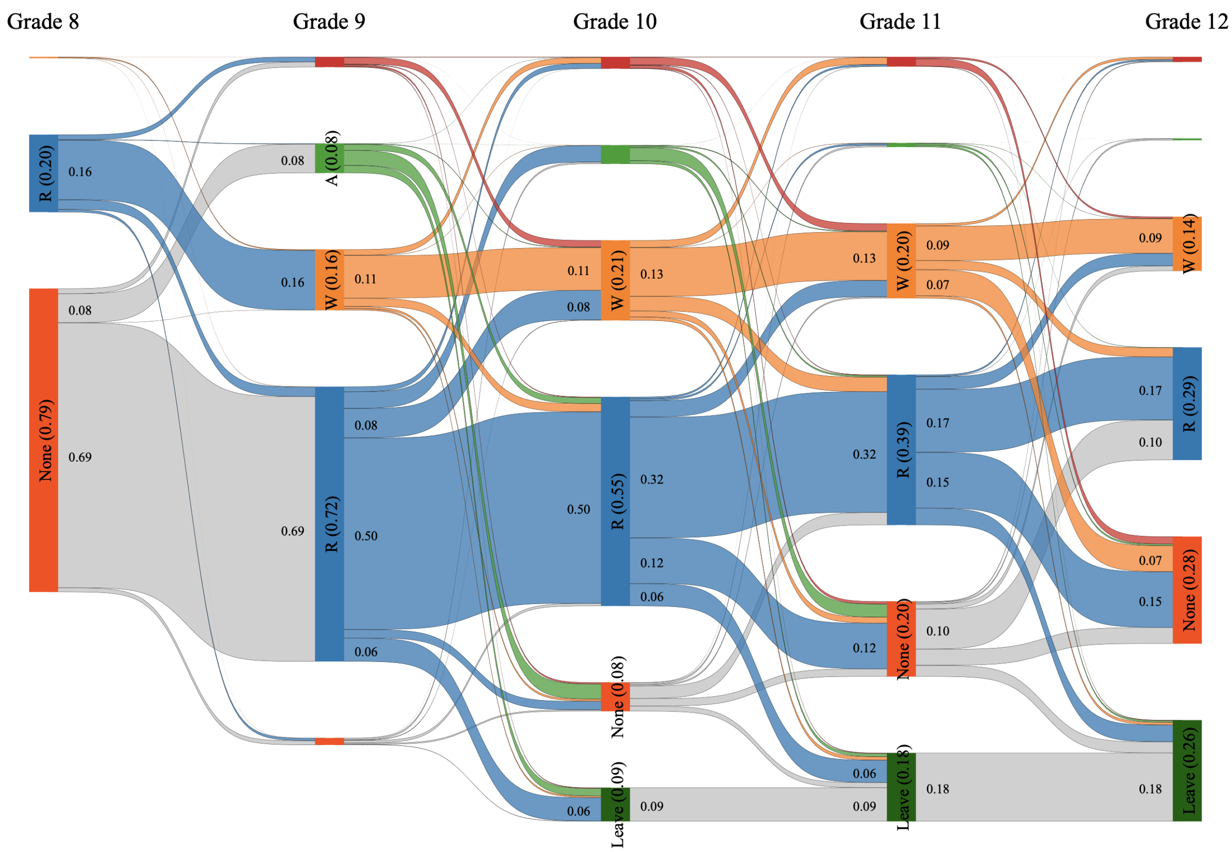
Note: The figure shows the distribution of GPAs for students after grades 9, 10, 11, and 12. GPA data for students after completing grades 9-12. Mean GPA and standard deviation are provided for each grade level. The distributions appear to be roughly normal, with slight changes in shape and central tendency across grade levels.
Data source: Administrative data from a large urban district in Texas

Course-taking Patterns: Figure 3 visualizes the math course progression of students from Grade 8 through Grade 12, highlighting various pathways that students take throughout their high school years. The diagram categorizes students into different groups based on their course-taking behavior: Regular (R), Weighted (W), Accelerated (A), None (no rank-eligible math courses), and Leave (students who exit the school system). The flow lines between these categories represent transitions between courses as students progress through high school, with the thickness of the lines corresponding to the proportion of students making each transition.

In Grade 8, a majority of students (79%) do not take any rank-eligible math courses. Most of these students transition into the regular math option (Algebra I) in Grade 9, while a small proportion pursue the accelerated math option (Algebra I + Geometry) in Grade 9. On the other hand, 20% of the students take Algebra I in Grade 8, and a majority of them subsequently enroll in the weighted math option (Geometry Pre-AP) in Grade 9. This early acceleration in middle school sets the foundation for these students to continue in more advanced math courses in high school.

In Grade 9, 72% of students take the regular math option, while 16% opt for weighted courses, and 8% pursue accelerated math options. Students enrolled in regular courses in Grade 9 tend to persist in the regular track in subsequent grades, although there is some movement into other tracks over time. The proportion of students taking regular math courses decreases gradually as students move to higher grades, while there is a notable increase in the proportion of students opting not to take any rank-eligible math courses as they approach the end of high school. Similarly, students who take the weighted option in Grade 9 are more likely to continue with the weighted course progression in the following years, but there is some leakage into regular and other course options.

Figure 3. Student Progression and Academic Pathway Flows Across Grades 8-12



Note: This Sankey diagram illustrates the academic pathways of students from Grade 8 through Grade 12. The width of each flow represents the proportion of students following that particular path. The main categories tracked appear are R ("Regular" courses), W ("Weighted" courses), A ("Accelerated" courses), "None" (which might indicate students not enrolled in specific rank-eligible courses), "Leave" (which indicates leaving school). The visualization demonstrates how students transition between these academic categories as they progress through high school grades.
Data source: Administrative data from a large urban district in Texas

Course-taking and Academic Standing Table 3 presents the summary statistics conditional on the math course chosen, where we distinguish math course options into five categories: leaving school, staying but not taking any (rank-eligible) math course, taking a regular (unweighted) math course (e.g., Geometry after completing Algebra I), taking a weighted math course (e.g., Pre-AP Geometry after Algebra I), or taking an accelerated course (e.g., taking both Geometry and Algebra II after Algebra I). Panel A displays the beginning-of-year GPA, rank, and math progression statistics. Students enrolled in weighted or accelerated math courses have, on average, higher GPAs and ranks at the start of the year compared to those taking regular courses. For instance, the average GPA for students taking weighted or accelerated courses is 86.92 and 86.17, respectively, compared to 84.73 for those in regular courses. Similarly, the average rank percentile is 0.61 and 0.56 for students in weighted and accelerated courses, respectively, compared to 0.49 for those in regular courses.

We also examine whether students are on the recommended math progression, which typically involves completing Algebra I in ninth grade, Geometry in tenth grade, Algebra II in eleventh grade, and an advanced course like Pre-Calculus in twelfth grade. Students who have accelerated their math progression (e.g., by completing Algebra I in middle school) are more likely to take weighted courses. On average, 71% of students enrolled in weighted courses have accelerated previously, compared to only 6% of those in regular courses.

Panel B provides residualized GPA and rank statistics, controlling for past academic performance and demographic characteristics. The negative coefficients indicate that students with lower residualized GPA and rank are more likely to leave school or take fewer rank-eligible courses, while those with higher residuals are more likely to enroll in weighted courses.

Panel C summarizes the end-of-year GPA and rank for each group, showing that weighted and accelerated course-takers continue to outperform their peers by the end of the year, with average GPAs of 86.49 and 85.50, respectively, compared to 84.81 for regular course-takers. Similarly, those who enroll in the weighted or accelerated courses have a higher rank – 0.58 and 0.52, respectively – at the end of the year compared to those who enroll in the regular math course – 0.49.

Overall, the table highlights that students who take weighted or accelerated courses have stronger academic backgrounds and are more likely to have deviated from the standard math progression. At the same time, these students go on to have a higher academic performance in the future as well. The results suggest that early acceleration and strong academic performance are correlated with more advanced course-taking in high school. These patterns may also reflect learning – students upon learning about higher-than-expected GPA and rank may increase their enrollment in rigorous coursework.

5 Model

5.1 Overview

Motivated by the descriptive patterns in the data, we now turn to our model of high school course decisions. Individuals in each period from $t = 1$ to $t = T$ make a decision regarding their mathematics course-taking. A student's options include whether to leave school, stay in school but not

Table 3. Past and Future academic characteristics by course choice

	(1) Leave	(2) None	(3) Regular	(4) Weighted	(5) Accelerated	(6) Total
Panel A: Beginning-of-Year GPA, Rank, and Math Progression						
GPA	82.81	84.61	84.73	86.92	86.17	85.03
	(4.30)	(4.09)	(3.98)	(4.10)	(5.10)	(4.28)
Rank	0.36	0.46	0.49	0.61	0.56	0.50
	(0.29)	(0.29)	(0.28)	(0.27)	(0.30)	(0.29)
Above the recommended level	0.17	0.18	0.06	0.71	0.21	0.20
	(0.38)	(0.38)	(0.24)	(0.45)	(0.40)	(0.40)
Panel B: Beginning-of-Year Residualized GPA and Rank						
GPA	-3.37^{***}	-2.16^{***}	-1.29^{***}		-0.15^{***}	
	(0.03)	(0.02)	(0.02)		(0.04)	
Rank	-3.37^{***}	-2.16^{***}	-1.29^{***}		-0.15^{***}	
	(0.03)	(0.02)	(0.02)		(0.04)	
Panel A: End-of-Year GPA and Rank						
GPA	-	84.63	84.81	86.49	85.50	85.17
	-	(3.96)	(4.04)	(4.05)	(4.89)	(4.15)
Rank	-	0.44	0.49	0.58	0.52	0.50
	-	(0.29)	(0.28)	(0.28)	(0.30)	(0.29)
Obs	37,683	145,652	230,602	75,380	25,384	514,701

Note: This table reports summary statistics for the data that is used to estimate our structural model. Standard deviations are listed directly below the mean (in parentheses) for each entry. This is student-grade-level pooled data.
Data source: Administrative data from a large urban district in Texas

take a (rank-eligible) mathematics course, take the regular mathematics course, take the weighted mathematics course, or take the accelerated mathematics course.

Students have imperfect information about their mathematics ability, which we denote by A_i. We assume that A_i is also unobserved to the econometrician and normally distributed, with mean zero and variance $\sigma^2_{g,a}$. Additionally, students have imperfect information about the average ability of their cohort, which we denote by R_{-i}. We assume that R_{-i} is also unobserved to the econometrician and normally distributed, with mean zero and variance $\sigma^2_{r,a}$. Beyond these individual and relative abilities that are initially unknown to the students, we also allow for unobserved (to the econometrician only) heterogeneity.

Students update their beliefs about A_i by receiving GPA signals that depend on their choices. These signals reveal information regarding their ability. Similarly, students update their beliefs about R_{-i} by receiving rank signals. These signals reveal information regarding the average ability of their cohort.

Students are assumed to be forward-looking and choose the sequence of actions yielding the highest value of expected lifetime utility. Hence, when making their course decisions, individuals consider the option value associated with the new information acquired on different choice paths.

We now detail the main elements of the model. We first discuss the grade and rank production functions. We then describe how individuals update their beliefs about their individual and relative abilities. Finally, we model the flow payoffs and the optimization problem the individuals face. Discussions of model identification and estimation are deferred to Sections 6 and 7, respectively.

5.2 Timing:

The model is designed to capture the progression of high school students through different grade levels, denoted by $t = 1, \ldots, T$, where $T = 4$. Each grade level represents a specific point in time during a student's academic journey. The timing of the grade levels is as follows: $t = 1$ corresponds to the first year of high school (aka ninth grade or freshman year), $t = 2$ corresponds to the second year of high school (aka tenth grade or sophomore year), $t = 3$ corresponds to the third year of high school (aka eleventh grade or junior year), and $t = 4$ corresponds to the fourth year of high school (aka twelfth grade or senior year).

5.3 Choice Variable:

In each grade level $t \in 1, \ldots, T$, student i makes a decision regarding their math course-taking, denoted by $a_{it} \in 0, \ldots, 4$ corresponding to leave, no math course, recommended course, weighted course, accelerated course. In other words, they can leave school, stay in school but not take a (rank-eligible) mathematics course, take the regular mathematics course, take the weighted mathematics course, or take the accelerated mathematics course. Mathematics in high school is taken in progression. We define the math pre-requisite level, represented by $\ell_{it} \in \mathcal{L}$, which indicates the highest math course student i has taken before reaching grade level t, and therefore determines their eligibility to take certain courses in the subsequent grade levels. The set $\mathcal{L}$ consists of six elements: $0, 1, 2, 3, 4, 5$ corresponding to the pre-requisite levels of none, Algebra I, Geometry, Algebra II, Pre-Calculus, Calculus, respectively.

For example, if $\ell_{it} = 0$, it means student i has not taken any math course before reaching grade level t, and they are eligible to take Algebra I in the current grade level (t), or in future grade levels. If $\ell_{it} = 1$, it indicates that the student has completed Algebra I and is now eligible to take Geometry. Similarly, if $\ell_{it} = 2$, the student has completed Geometry and can proceed to Algebra II, and so on. Hence, based on their pre-requisite level ℓ_{it}, student i makes a math course decision $a_{it} \in \mathcal{A}\ell$ where $\mathcal{A}\ell$ represents the choice set available to a student if their pre-requisite level in grade t is ℓ, i.e., $\ell_{it} = \ell$. The choice sets are summarized in Table 4. Alternatives 1 and 2 mean leave school and stay in school but not take any (rank-eligible) math course, irrespective of the math pre-requisite level. Someone who has not taken any high school mathematics course yet ($\ell = 1$) can choose from: Leave, No math course, Algebra I, Algebra I Pre-AP, and Geometry. Someone who has completed Algebra I can choose from: Leave, No math course, Geometry, Geometry Pre-AP, and Algebra II. Someone who has completed Geometry can choose from: Leave, No course, Algebra II, Algebra II Pre-AP, and Pre-Calculus. Someone who has completed Algebra II can choose from: Leave, No course, Pre-Calculus, Pre-Calculus Pre-AP, and Calculus AP. Note that Calculus is only available in the weighted (AP) version. Someone who has completed Pre-Calculus can choose from: Leave, No course, and Calculus AP. Note that there is no math course more advanced than Calculus.

Table 4. State-specific Choice set

ℓ \ $\mathcal{A}_\ell$	Regular	Weighted	Accelerated
None	Algebra I	Algebra I (Pre-AP)	Geometry
Algebra I	Geometry	Geometry (Pre-AP)	Algebra II
Geometry	Algebra II	Algebra II (Pre-AP)	Pre-Calculus
Algebra II	Pre-Calculus	Pre-Calculus (Pre-AP)	-
Pre-Calculus	-	Calculus (AP)	-

Note: ℓ refers to the math pre-requisite level attained before making the course choice and $\mathcal{A}_\ell$ refers to the set of alternatives available to someone who is at pre-requisite level ℓ. Alternatives 1 and 2 mean leave school and stay in school but not take any math course, irrespective of the math pre-requisite level, so they are omitted from the table to save space.

5.4 Flow Utility

The flow payoff for each math course decision is influenced by three primary factors. First, a student's incentives are shaped by their observed, time-varying ability as measured by their class rank, denoted as $r_{it} \in [0, 1]$. This rank is based on the student's high school class percentile, where a higher value indicates a better rank, reflecting the student's relative performance compared to peers in the cohorts (i.e., individuals who graduate from the same school in the same year).

Second, there is a time-varying, unobserved utility component specific to course a, denoted as $\epsilon_{it}(a)$. This component captures individual-level idiosyncratic preferences for each course that vary over time. For example, changes in personal interests or external circumstances could influence a student's inclination toward particular math courses at different stages.

Third, there is an unobserved, time-invariant utility component, $\beta_k(a)$, which depends on their unobserved type k. This component captures persistent preferences for a given course option based on the student's type, which could reflect characteristics such as intrinsic ability, motivation, or external constraints that do not change over time. This unobserved heterogeneity allows for variation in students' preferences that are not captured by observable characteristics and persist across different grade levels.

In addition to these three factors, utility also depends on whether the student has completed the recommended math course before entering that grade level, represented as $\ell_{it} > \bar{\ell}_t$. Here, ℓ_{it} indicates the highest math course completed by student i before grade t and $\bar{\ell}_t$ is the recommended level of math proficiency for grade t. If a student is above this recommended level, they may derive additional utility from feeling more prepared or confident in tackling advanced coursework. Finally, the student's demographic information X_i can also influence their utility from different course options.

The utility from leaving school, i.e., $a = 0$, is normalized to zero. In grade level $t \in \{1, \ldots, T\}$, the flow payoff for student i from choice $a \in \{1, 2, 3, 4\}$ is given by:

$$u_t(a, \ell_{it}, r_{it}, \epsilon_{it}; \beta) = \beta_{a,0} + \beta_{a,\ell}(\ell_{it} > \bar{\ell}_t) + \beta_{a,r,1} r_{it} + \beta_{a,r,2} r_{it}^2 + \beta_{a,x} X_i + \sum_k \beta_{k,a}(k_i = k) + \epsilon_{it}(a). \quad (1)$$

Here, $\beta_{a,0}$ represents the base utility level of choosing course a while $\beta_{a,\ell}$ captures how this base utility differs for those above the recommended math level. This specification allows the utility to vary by whether the student has accelerated beyond the standard math progression. The recommended levels are: no math course required before entering Grade 9, Algebra I before Grade 10, Geometry before Grade 11, and Algebra II before Grade 12, i.e., $\bar{\ell}_t = t - 1$. Additionally, the terms $\beta_{a,r,1}$ and $\beta_{a,r,2}$ capture the linear and quadratic effects of class rank on utility. The inclusion of both terms allows the relationship between class rank and utility to be nonlinear, reflecting potential diminishing or increasing returns to rank as students move up in the distribution.

5.5 Terminal Utility

At the end of high school, student has no more math course choices to make but they now have their final class rank r_{iT+1} that shapes their terminal utility:

$$u_{T+1}(r_{iT+1}; \beta) = \beta_{r,1} r_{iT+1} + \beta_{r,2} r_{iT+1}^2. \tag{2}$$

5.6 Rank Production and Learning

In this subsection, we delve into the process by which students form and update their beliefs about their academic rank within their cohort. We introduce a model that captures the relationship between a student's GPA and their rank, while accounting for the uncertainty surrounding the average ability of their peers. By receiving noisy signals in the form of realized GPA and rank outcomes, students engage in Bayesian updating to refine their beliefs about the cohort's ability distribution. This learning process generates endogenous heterogeneity in perceived ability, even among students with identical initial priors, highlighting the role of heterogeneous experiences in shaping academic decisions.

Central to a student's uncertainty about their future rank is their lack of knowledge about the average ability of their cohort. While students understand that their rank depends on their own GPA, they are also aware that it is influenced by the GPA distribution of their peers, which is unobserved. To capture this relationship, we propose the following model:

$$r_{it} = \Phi(\alpha_g g_{it} - R_{-i} + \eta_{r,it}) \tag{3}$$

where r_{it} represents the rank of student i in grade level t, g_{it} is their GPA before entering grade level t, and

where $R_{-i} \sim \mathcal{N}(\tilde{\mu}_r, \tilde{\sigma}_r^2)$ denotes the average ability of the cohort relative to student i. The cohort's ability distribution is assumed to be normal with mean $\tilde{\mu}_r$ and variance $\tilde{\sigma}_r^2$, both of which are unknown to the student and the econometrician. The parameter α_g captures the influence of the student's GPA on their rank, while $\eta_{r,it} sim \mathcal{N}(0, \gamma_r^2)$ is an idiosyncratic error term. The function $\Phi(\cdot)$ represents the cumulative distribution function (cdf) of the standard normal distribution.

But like we said before, the student does not know the true distribution $\tilde{\mu}_r, \tilde{\sigma}_r^2$. Students are unsure about the cohort's relative ability. We assume that students are rational and update their beliefs in a Bayesian fashion. Students receive rank signals but before they receive any information We assume that student has a prior belief about the cohort ability which is normally distributed with mean zero and variance σ_r^2.

In the absence of perfect information about the cohort's ability distribution, students are assumed to hold prior beliefs and update them rationally as they receive new information. Initially, students' prior beliefs about the cohort's ability are normally distributed with mean zero and variance σ_r^2. Throughout their high school years, students observe their realized GPA and rank outcomes, which serve as noisy signals about the true cohort ability. The signal received by student i in grade level t is given by:

$$s_{r,it} \equiv R_{-i} - \eta_{r,it} = \alpha_g g_{it} - \Phi^{-1}(r_{it}). \tag{4}$$

This signal is the difference between the student's realized GPA and the inverse of the standard normal cdf evaluated at their realized rank. Intuitively, it captures the discrepancy between the student's actual rank and the rank they would expect given their GPA, providing information about the cohort's ability.

Students incorporate these signals into their beliefs using Bayesian updating. The updated beliefs about the cohort's ability distribution are characterized by the following equations:

$$\sigma_{r,t+1}^2 = \left(\frac{1}{\sigma_{r,t}^2} + \frac{1}{\gamma_r^2}\right)^{-1} \tag{5a}$$

$$\mu_{r,t+1} = \left(\frac{1}{\sigma_{r,t}^2} + \frac{1}{\gamma_r^2}\right)^{-1} \left(\frac{1}{\sigma_{r,t}^2}\mu_{r,it} + \frac{1}{\gamma_r^2} s_{r,it}\right) \tag{5b}$$

where $\sigma_{r,t}^2$ represents the perception error variance, which decreases as the student accumulates more information. The magnitude of the update to the mean belief, $\mu_{r,t}$, depends on the signal's accuracy, with more precise signals (i.e., smaller γ_r^2) leading to larger updates. Notably, when a student's realized rank is lower than their expected rank, their belief about the mean of the cohort's ability distribution shifts upward.

A key insight from this learning model is that the signals received by students, $s_{r,it}$, are inherently heterogeneous due to the randomness in the realized GPA and rank outcomes. As a result, even students who start with identical prior beliefs about the cohort's ability will develop different perceived abilities over time. This endogenous heterogeneity in beliefs, driven by heterogeneous experiences, offers a compelling explanation for the observed variation in course choices that goes beyond the traditional assumption of inherent preference differences. By incorporating rank production and learning into our model, we capture the dynamic process by which students form and update their beliefs about their academic standing within their cohort.

5.7 GPA Production and Learning

We now turn our attention to the evolution of a student's GPA and the role of learning in shaping their beliefs about their own ability. We propose a model that captures the relationship between a student's course decisions, past academic performance, and unobserved heterogeneity in determining their GPA. Alongside this production process, we introduce a Bayesian learning mechanism that allows students to update their beliefs about their own ability based on the noisy signals they receive in the form of realized GPA outcomes.

We model the evolution of a student's GPA as a function of their course decision (a_{it}), the highest math course taken before the current grade level (ℓ_{it}), their GPA in the previous grade level (g_{it}), and an idiosyncratic shock ($\eta_{it}(a)$). Additionally, we incorporate unobserved heterogeneity in the form of discrete types (k_i) that are known to the student but unobserved by the econometrician, as well as a time-invariant ability component (A_i) that is unobserved to both the student and the econometrician. Formally, the GPA production function is given by:

$$\begin{aligned} g_{it+1}(a, \ell_{it}, g_{it}, \eta_{it}; \alpha) &= \alpha_{a,0} + \alpha_{a,\ell}(\ell_{it} > \bar{\ell}_t) + \alpha_{a,g} g_{it} + \sum_k \alpha_{a,k}(k_i = k) + A_i + \eta_{g,it}(a) \\ &= f_g\left(a, \ell_{it}, g_{it}, k_i; \theta_g\right) + A_i(a) + \eta_{g,it}(a) \end{aligned} \tag{6}$$

where $A_i(a) \sim \mathcal{N}(\tilde{\mu}_{g,a}, \tilde{\sigma}_{g,a})$ represents the time-invariant ability of student i that is unobserved to the student and the econometrician. GPA also depends on $\beta_{a,k}$ that is the unobserved heterogeneity for student of type k that is known to the student but not the econometrician. We assume that the idiosyncratic shocks, $\eta_{g,it}(a)$, are mutually independent and distributed $\mathcal{N}(0, \gamma_g(a)^2)$, and are also independent from the other state variables. This specification allows for level shifts in GPA based on course alternatives and grade level through $\alpha_{a,0}$. Through $\alpha_{a,g}$, it allows for variation in student's GPA at the time of entering grade level $t+1$ by their GPA at the time of entering grade level t.

As with the cohort's ability distribution, students do not have perfect knowledge of their own ability $A_i(a)$ and the associated parameters $\tilde{\mu}_{g,a}(a)$ and $\tilde{\sigma}_{g,a}(a)$. Instead, they hold prior beliefs about their ability, which are assumed to be normally distributed with mean zero and variance $\sigma^2_{g,a}(a)$. As students progress through high school, they observe their realized GPA outcomes, which serve as noisy signals about their true ability. The signal received by a type-k student i in grade level t from course decision a is given by:

$$s_{g,it}(a, k) \equiv A_i(a) + \eta_{g,it-1}(a) = g_{it} - f_g\left(a, \ell_{it-1}, g_{it-1}, k; \theta_g\right). \tag{7}$$

This signal represents the difference between the student's realized GPA and their expected GPA based on their observable characteristics and course decision. Students incorporate these signals into their beliefs using Bayesian updating, with the updated beliefs characterized by the following equations:

$$\sigma^2_{g,it+1}(a) = \left(\frac{1}{\sigma^2_g(a)} + \frac{\sum_{\tau=0}^{t}(a_{i\tau} = a)}{\gamma^2_g(a)} \right)^{-1} \tag{8a}$$

$$\mu_{g,it+1}(a,k) = \left(\frac{1}{\sigma^2_g(a)} + \frac{\sum_{\tau=0}^{t}(a_{it} = a)}{\gamma^2_g(a)} \right)^{-1} \left(\frac{1}{\sigma^2_{gt}(a)} \mu_g(a) + \frac{1}{\gamma^2_g(a)} \sum_{\tau=0}^{t+1} s_{g,i\tau}(a,k)(a_{i\tau} = a) \right) \tag{8b}$$

where $\sigma^2_{g,t}(a)$ represents the perception error variance, which decreases as the student accumulates more signals. The magnitude of the update to the mean belief, $\mu_{g,it}(a,k)$, depends on the precision of the signals, with more precise signals (i.e., smaller $\gamma^2_g(a)$) leading to larger updates. When a student's realized GPA is higher than their expected GPA, their belief about their own ability shifts upward.

As with the learning process for cohort ability, the heterogeneity in the signals received by students, $s_{g,it}(a,k)$, leads to endogenous differences in perceived ability across individuals, even when they start with the same priors and make the same course decisions. This feature of the model highlights the importance of heterogeneous experiences in shaping students' beliefs about their own ability, rather than relying solely on inherent preference differences to explain variation in course choices.

The GPA production and learning mechanism, in conjunction with the rank production and learning process described in the previous subsection, provides a comprehensive framework for understanding how students form and update their beliefs about their academic abilities and standing within their cohort. By incorporating these learning processes into our model, we can better capture the dynamic nature of student decision-making and the role of uncertainty and information in shaping educational outcomes.

5.8 The Optimization Problem

We now consider the dynamic optimization problem faced by students as they make their course decisions throughout high school. We assume that students are forward-looking and choose the sequence of courses that maximizes the expected present value of their lifetime utility, taking into account the uncertainty surrounding their future preferences and ability signals.

Formally, the student's objective is to choose a sequence of course decisions $(a_{it})_{t=1\ldots T}$ that maximizes the discounted sum of expected payoffs:

$$E\left[\sum_{t=1}^{T} \delta^{t-1} \sum_{a} \left(u_t(a, \ell_{it}, r_{it}; \beta) + \epsilon_{it}(a) \right) 1\left\{ a_{it} = a \right\} \right] \tag{9}$$

where $\delta \in (0,1)$ is the discount factor, $u_t(\cdot)$ represents the flow utility derived from course decision a in grade level t, and $\epsilon_{it}(a)$ captures idiosyncratic preference shocks. The expectation is taken with respect to the distribution of future preference shocks, and ability signals, conditional on the student's information set at each decision point.

To characterize the optimal course sequence, we introduce the ex-ante value function $V_t(\ell_{it}, r_{it}, k)$, which represents the expected discounted sum of current and future payoffs at the beginning of grade level t, before the realization of the idiosyncratic preference shock. The conditional value function $v_t(a, \ell_{it}, r_{it}, k; \beta)$, which represents the value of choosing course a in grade level t given the state variables and the realized preference shock, can be expressed as:

$$v_t\left(a, \ell_{it}, r_{it}, k; \beta\right) = u_t\left(a, \ell_{it}, r_{it}, k; \beta\right) + \delta E_t\left[V_{t+1}\left(\ell', r', k\right) \mid \ell_{it}, r_{it}, a_{it} = a\right] \tag{10}$$

where $E_t[\cdot \mid \cdot]$ denotes the expectation conditional on the student's information set at the beginning of grade level t, which includes the sequence of ability signals received up to period $t-1$.

Assuming that the preference shocks $\epsilon_{it}(a)$ are independently and identically distributed according to a Type 1 extreme value distribution, the conditional value function for $t < T$ can be written as the following weighted log-sum formula:

$$v_t\left(a, \ell_{it}, r_{it}, k; \beta\right) = u_t\left(a, \ell_{it}, r_{it}, k; \beta\right) + \delta E\left[\ln \sum_{a'} \left(\exp(v(a', \ell', r', k; \beta)\right) \mid \ell_{it}, r_{it}, a_{it} = a, k\right] + \delta\Gamma \text{ if } t < T \tag{11a}$$

$$= u_t\left(a, \ell_{it}, r_{it}, k; \beta\right) + \delta E\left[u_{t+1}(r'; \beta) \mid \ell_{it}, r_{it}, a_{it} = a, k\right] \text{ if } t = T \tag{11b}$$

where Γ denotes Euler's constant.

To estimate the model, we leverage the conditional choice probability (ccp) inversion method proposed by Hotz and Miller (1993). By exploiting the fact that leaving school ($a = 1$) is a terminal action, we can rewrite the conditional value function as:

$$v_t\left(a, \ell_{it}, r_{it}, p_{it+1}, k; \beta\right) = u_t\left(a, \ell_{it}, r_{it}, k; \beta\right) + \delta E\left[\ln \sum_{a'} \left(\exp(\log p_{it+1}(a', \ell', r') - \log p_{it+1}(1, \ell', r')\right) \mid \ell_{it}, r_{it}, a_{it} = a \right. \tag{12a}$$

$$= u_t\left(a, \ell_{it}, r_{it}, k; \beta\right) + \delta E\left[u_{t+1}(r'; \beta) \mid \ell_{it}, r_{it}, a_{it} = a, k\right] \text{ if } t = T \tag{12b}$$

The ccp inversion approach allows us to express the continuation value in terms of future choice probabilities, which can be estimated from the data. This reformulation simplifies the dynamic optimization problem and facilitates the estimation of the structural parameters governing students' preferences and beliefs.

6 Identification

In this section, we discuss the identification of the key components of our dynamic discrete choice model, including unobserved heterogeneity, conditional choice probabilities, conditional value func-

Table 5. Mathematical Notations

Symbol	Description	Main equations of reference
a_{it}	Course choice of student i in period t	(1) , (6)
ℓ_{it}	Math prerequisite level of student i in period t	(1), (6)
r_{it}	Class rank of student i in period t	(1), (3), (9)
$\epsilon_{it}(a)$	Idiosyncratic preference shock for course a	(1), (9)
β	Parameters in utility function	(1), (9)
g_{it}	GPA of student i in period t	(3), (6)
R_{-i}	Average ability of cohort relative to student i	(3), (4)
$\eta_{r,it}$	Idiosyncratic error in rank production	(3), (4)
α_g	Parameter in rank production function	(3), (4)
σ_r^2	Variance of prior belief about cohort ability	(5)
γ_r^2	Variance of rank signal noise	(5)
$\mu_{r,t}$	Mean of posterior belief about cohort ability	(5)
$A_i(a)$	Time-invariant ability of student i for course a	(6)
$\eta_{g,it}(a)$	Idiosyncratic shock in GPA production	(6)
$\sigma_g^2(a)$	Variance of prior belief about individual ability	(8)
$\gamma_g^2(a)$	Variance of GPA signal noise	(8)
$\mu_{g,it}(a)$	Mean of posterior belief about individual ability	(8)
V_t	Value function	(10), (11)
v_t	Conditional value function	(10), (11), (12)
p_{it}	Conditional choice probability	(12)
$q_i(k)$	Posterior probability of being type k	(22)
$\pi_{k\|x}$	Type probability conditional on initial state	(23)

tions, flow utilities, GPA and rank production functions, and unobserved abilities. We highlight the assumptions and variations in the data that allow us to separately identify these components and estimate the structural parameters of interest.

6.1 Unobserved Heterogeneity

One of the primary challenges in identifying the parameters of our model is the presence of unobserved preferences that are known to the students but not to the researcher. These unobserved preferences may be correlated over time, leading to potential biases in the estimated effects of observed characteristics on students' choices and outcomes. To address this issue, we incorporate permanent unobserved heterogeneity following the approach of Keane and Wolpin (1997). Specifically, we allow for K distinct unobserved types of students, with each type characterized by a unique set of type-specific components that shift the intercepts of the GPA transition and flow utility functions.

The identification of the distribution of unobserved heterogeneity relies on the dynamic nature of our model. By comparing the choices of observationally equivalent students over time, we can infer the presence of different unobserved types (Arcidiacono et al., 2025). Intuitively, if some students consistently make choices that deviate from what we would expect based on their observed characteristics and past outcomes, this suggests the presence of unobserved factors that drive their behavior. For example, if a student with good grades chooses to drop out of school, this may indicate that they belong to a type with low intrinsic motivation or low utility from attending school. Similarly, if a student chooses a course in which they have lower predicted ability compared to other courses, this may reflect a high idiosyncratic preference for that particular course.

The identification of unobserved types also relies on the variation in outcomes across observationally equivalent students. Among students with the same observed characteristics and choice histories, those who perform better in a given course are more likely to belong to a type with higher unobserved ability or match quality for that course. By comparing the choice probabilities and outcome distributions of observationally equivalent students, we can estimate the proportions of each unobserved type in the population and their associated type-specific parameters.

A crucial assumption for the identification of unobserved heterogeneity is that the unobserved types are discrete, while the unobserved abilities are continuous and normally distributed. This assumption allows us to separately identify the effects of permanent type-specific preferences from the effects of unobserved abilities that are initially unknown to the students. Intuitively, the discreteness of the unobserved types generates "lumpy" variation in choices and outcomes that cannot be fully explained by the smooth variation in unobserved abilities. The observed variation in choices and outcomes among observationally equivalent students is thus more likely to reflect differences in unobserved types rather than differences in unobserved abilities.

However, the identification of unobserved heterogeneity is complicated by the fact that the unobserved types and abilities may be correlated, conditional on the observed outcomes. For example, the observed GPA of a student who chooses course a depends not only on their course-specific ability but also on their unobserved type, which may affect both their choice of a and their performance in a. To disentangle the effects of unobserved types and abilities, we rely on the functional form assumptions of discrete types and normally distributed abilities, as well as the exclusion restriction that the flow utilities depend only on the observed rank and not on the unobserved abilities.

6.2 Conditional Choice Probabilities

Conditional on the observed state variables, which include the observed characteristics and past choices of students, the choice probabilities can be expressed as a finite mixture of type-specific conditional choice probabilities, where the mixture weights correspond to the type proportions identified in the previous step. The identification of type-specific conditional choice probabilities in dynamic discrete choice models with unobserved heterogeneity has been studied by Kasahara and Shimotsu (2009) and Hu and Shum (2012). A key requirement for identification is that the observed state variables generate sufficient variation in choices across types, such that the matrix of conditional choice probabilities is full rank. Intuitively, this means that the observed variation in choices across states must be "rich enough" to distinguish between the different unobserved types.

6.3 Conditional Value Functions

Once the type-specific conditional choice probabilities have been identified, the conditional value functions associated with each choice alternative can be identified using standard arguments from the dynamic discrete choice literature (Hotz and Miller, 1993; Arcidiacono and Miller, 2011). The key requirement is the assumption that the idiosyncratic preference shocks are distributed according to a Type 1 extreme value distribution, which allows us to invert the conditional choice probabilities to obtain the conditional value functions. The scale and location of the conditional value functions are not separately identified, so we normalize the value function for one reference alternative (e.g., dropping out of school) to zero.

To simplify the computation of the conditional value functions, we follow the approach of Arcidiacono and Miller (2011) and Arcidiacono and Miller (2019) and re-express the future value terms in the Bellman equation in terms of a few period-ahead conditional choice probabilities and flow utilities, rather than solving the full dynamic programming problem. This is possible due to the linearity of the flow utilities in the parameters and the additive separability of the preference shocks, which allows us to write the expected maximum of the conditional value functions as a function of the logarithms of the conditional choice probabilities, also known as the "log-sum formula".

6.4 Flow Utilities

The flow utility parameters are identified from the variation in choices across students with different observed characteristics and choice histories. The key identifying assumption is that the observed variation in choices reflects differences in the flow utilities associated with each choice alternative, conditional on the expected future value terms that depend on the ability beliefs and the transition probabilities of the observed states. The flow utility parameters are estimated by matching the observed choice probabilities to the choice probabilities predicted by the model, which are functions of the flow utilities and the conditional value functions.

To ensure that the scale of the flow utilities is identified, we normalize the flow utility of one reference alternative (e.g., dropping out of school) to zero and fix the scale of the Type 1 extreme value shocks to 1. Given these normalizations, the remaining flow utility parameters can be identified from the differences in the log odds of choosing each alternative relative to the reference alternative, for different values of the observed states. Intuitively, if students with certain observed characteristics or choice histories are more likely to choose a particular course relative to the reference alternative, this suggests that the course has higher flow utility for those students.

6.5 GPA and Rank Production Functions

The parameters of the GPA and rank production functions are identified from the variation in observed GPA and rank outcomes across students with different observed characteristics, choice histories, and unobserved types. A key challenge in identifying these parameters is the potential for sample selection bias, since the GPA and rank outcomes for each course are only observed for students who actually choose to take that course. This means that the observed variation in outcomes may reflect not only the causal effects of the observed characteristics and choices on performance but also the unobserved preferences and abilities that drive course selection.

To address this issue, we rely on two key assumptions. First, we assume that the unobserved types are independent of the observed characteristics and past choices, conditional on the student's unobserved abilities. This allows us to express the expected outcome for each course as a function of the student's observed characteristics, past choices, unobserved abilities, and a course-specific "match effect" that depends on the student's unobserved type. Second, we assume that the unobserved abilities enter the GPA and rank production functions additively and are independent of the other inputs, conditional on the unobserved type. This allows us to write the expected outcomes as the sum of a type-specific component that depends on the observed inputs and a type-specific mean ability term.

Given these assumptions, the parameters of the GPA and rank production functions can be identified from the variation in outcomes within and between unobserved types, controlling for the observed characteristics and past choices of students. The type-specific components are identified from the variation in outcomes across students with different observed inputs who are predicted to belong to the same unobserved type based on their choice histories. The mean ability terms are identified from the average differences in outcomes across unobserved types, controlling for the observed inputs. Intuitively, if students who are predicted to belong to a particular type based on their choices systematically perform better or worse than students with the same observed characteristics who are predicted to belong to other types, this suggests the presence of type-specific ability differences.

6.6 Unobserved Abilities

The means and variances of the unobserved ability distributions are identified from the observed persistence in GPA and rank outcomes over time, which depends on the true values of the unobserved abilities and the signal noise. Intuitively, if students who perform well in one period tend to perform well in subsequent periods, this suggests that their performance is driven by persistent unobserved abilities rather than idiosyncratic shocks. The ability variances are identified from the degree of persistence in outcomes, with higher variances implying greater persistence, all else equal.

The parameters of the learning process are identified from the observed changes in outcomes and choices over time as students accumulate more signals of their abilities. The speed of learning depends on the signal-to-noise ratios of the GPA and rank signals, which are identified from the relative magnitudes of the ability variances and the signal noise variances. Intuitively, if the observed variation in signals is mostly driven by noise rather than true ability differences, this implies that the signals are less informative and students will learn more slowly. The initial ability beliefs are identified from the variation in outcomes and choices in the first period, before any signals have been received.

Finally, the identification of the production function parameters and the learning parameters also relies on functional form assumptions, specifically the joint normality of the unobserved abilities and the normality of the signal noise. These assumptions allow us to tractably characterize the evolution of beliefs over time and to derive closed-form expressions for the choice probabilities and expected outcomes as functions of the parameters. While the functional form assumptions are not directly testable, they provide a parsimonious and computationally tractable way to model the learning process and are standard in the literature on Bayesian learning in dynamic choice models.

6.7 Summary

In summary, the identification of our model relies on a combination of functional form assumptions, exclusion restrictions, and variation in the observed choices and outcomes of students over time. There are five key identifying assumptions. First, the unobserved types are discrete and independent of the observed characteristics and choices, conditional on the unobserved abilities. Second, the unobserved abilities are normally distributed and enter the outcome equations additively, conditional on the unobserved types. Third, the idiosyncratic preference shocks are Type 1 extreme value distributed and independent of the other state variables. Fourth, the flow utilities depend only on the observed states and not on the unobserved abilities. Fifth, the choice-specific continuation values can be expressed as functions of a finite number of future choice probabilities and flow utilities.

Given these assumptions, the observed variation in choices and outcomes across students and over time provides sufficient information to identify the key structural parameters of the model, including the type probabilities, the type-specific choice probabilities and outcome parameters, the parameters of the learning process, and the means and variances of the unobserved ability distributions. These parameters can then be used to estimate the effects of counterfactual policies on students' choices, outcomes, and welfare.

7 Estimation

In this section, we describe the estimation method for the structural model parameters. We begin by presenting the estimation procedure without accounting for unobserved heterogeneity among students, as this simplifies the exposition and allows for a clearer understanding of the core steps. We then extend the estimation method to incorporate unobserved heterogeneity, which captures important differences across students that are not directly observable in the data.

7.1 GPA and Rank Production and Learning Parameters

We use Expectation-Maximization (EM) algorithms to estimate the parameters θ_g, $\sigma_g(a)$, θ_r and σ_r. The EM algorithm for the GPA production parameters alternates between two steps until reaching convergence:

E-step: We update the posterior ability distribution using all observed GPA data and the GPA production parameters from the previous iteration. Equation 8 provides the Bayesian updating formulas that update the posterior ability mean and variance. At each iteration h, we update the population variance of the ability distribution:

$$(\sigma_g(a)^{(h)})^2 = \frac{1}{N}\sum_{i=1}^{N}(\sigma_{g,i}(a)^{(h)})^2 + (\mu_{g,i}(a)^{(h)})^2 \tag{13}$$

where N represents the sample size, $\mu_{g,i}(a)^{(h)}$ denotes the posterior ability mean, and $\sigma_{g,i}(a)^{(h)}$ denotes the posterior ability standard deviation at the start of the E-step.

M-step: Using the posterior ability distribution from the E-step, we maximize the expected complete log-likelihood of the GPA data. Let $\varphi_{g,i}(\cdot)$ represent the pdf of the posterior ability distribution at iteration h. We maximize:

$$\begin{aligned}\mathbb{E}\left(L_{g,it}(a)^{(h)}\right) &= \int \log\left(\mathcal{L}_i(g_{it+1} \mid a, \ell_{it}, g_{it}, A_i(a))\right) \varphi_{g,i}^{(h)}(A_i(a))dA_i(a) \\ &= -\frac{1}{2}\log(2\pi\gamma_g(a)^2) - \frac{1}{2\gamma_g(a)^2}\left(\left(g_{it+1} - f_g(a, \ell_{it}, g_{it}; \theta_g^{(h)}) - \mu_{g,i}(a)^{(h)}\right)^2 + (\sigma_{g,i}(a)^{(h)})^2\right)\end{aligned} \tag{14}$$

where the normality assumptions on the idiosyncratic shocks and unobserved ability lead to the second equality. We update the parameters θ_g by solving:

$$\max_{\theta_g} \sum_{i,t,a} (a_{it} = a)\mathbb{E}\left(L_{g,it}(a)^{(h)}\right) \tag{15}$$

We follow a similar two-step EM procedure to estimate the rank production parameters:

E-step: Using observed rank data and rank production parameters from the previous iteration, we update the posterior cohort ability distribution. Equation 5 provides the Bayesian updating formulas that update the posterior ability mean and variance. At each iteration h, we update the population variance:

$$(\sigma_r^{(h)})^2 = \frac{1}{N}\sum_{i=1}^{N}(\sigma_{r,i}^{(h)})^2 + (\mu_{r,i}^{(h)})^2 \tag{16}$$

where $\mu_{r,i}^{(h)}$ and $\sigma_{r,i}^{(h)}$ represent the posterior cohort ability mean and standard deviation at the start of the E-step.

M-step: Using the posterior cohort ability distribution from the E-step, we maximize the expected complete log-likelihood of the rank data. Let $\varphi_{r,i}(\cdot)$ represent the pdf of the posterior ability distribution at iteration h. We maximize:

In the M-step, Given the posterior ability distribution obtained at the E-step, we maximize the expected complete log-likelihood of the rank data. Namely, at the M-step of each iteration h of the EM estimation, denoting by $\varphi_{r,i}(\cdot)$ the pdf of the posterior ability distribution computed at the E-step, we maximize the expected complete log-likelihood:

$$\begin{aligned}&\int \sum_t \log\left(\mathcal{L}_i(r_{it} \mid g_{it}, R_{-i})\right) \varphi_{r,i}(R_{-i})dR_{-i} \\ &= -\frac{1}{2}\log(2\pi\gamma_r^2) - \frac{1}{2\gamma_r^2}\left((\Phi^{-1}(r_{i\tau}) - \alpha_g g_{i\tau} + \mu_{r,i}^{(h)})^2 + (\sigma_{r,i}^{(h)})^2\right)\end{aligned} \tag{17}$$

We update the parameters (γ_r, α_g) by solving:

$$\max_{\gamma_r} \sum_{i,\tau} \left[-\frac{1}{2} \log(2\pi\gamma_r^2) - \frac{1}{2\gamma_r^2} \left(((\Phi^{-1}(r_{i\tau}) - \alpha_g g_{i\tau} - \mu_{r,i}^{(h)})^2 + (\sigma_{r,i}^{(h)})^2 \right) \right] \tag{18}$$

7.2 Utility Parameters

After estimating the GPA and rank production parameters, we focus on estimating the utility parameters θ_u that govern students' course choices.

To begin, we construct an initial guess of the conditional choice probabilities (CCPs) $p^{(1)}(a \mid \ell, r) = \Pr(a_{it} = a \mid \ell_{it} = \ell, r_{it} = r)$. We do this by grouping observations into bins based on discrete math level history $\ell_{it} \in \{0, \ldots, 5\}$ and discretized rank $r_{it} \in \{0.01, 0.02, \ldots, 0.99, 1.00\}$. For each (ℓ, r) bin, we compute the empirical frequency that students in that bin choose each course a. This gives us a non-parametric estimate of the CCPs.

Given $p^{(h)}$ and $\theta_u^{(h)}$, we solve for $\theta_u^{(h)}$ by maximizing:

$$\max_{\theta_u^{(h)}} \sum_{i,t,a} (a_{it} = a) L_{u,it}(a)^{(h)} \tag{19}$$

where

$$L_{u,it}(a)^{(h)} = \log \left[\frac{\exp\left(v_t \left(a, \ell_{it}, r_{it}, p_{it+1}^{(h)}; \theta_u^{(h)}, \hat{\theta}_g, \hat{\sigma}_g(a), \hat{\theta}_r, \hat{\sigma}_r \right) \right)}{\sum_a \exp\left(v_t \left(a, \ell_{it}, r_{it}, p_{it+1}^{(h)}; \theta_u^{(h)}, \hat{\theta}_g, \hat{\sigma}_g(a), \hat{\theta}_r, \hat{\sigma}_r \right) \right)} \right] \tag{20}$$

Then we update $p^{(h+1)}$

$$p_{it}(a, \ell, r)^{(h+1)} = \frac{\exp\left(v_t \left(a, \ell, r, p_{it+1}(\ell', r')^{(h)}; \theta_u^{(h+1)}, \hat{\theta}_g, \hat{\sigma}_g(a), \hat{\theta}_r, \hat{\sigma}_r \right) \right)}{\sum_a \exp\left(v_t \left(a, \ell, r, p_{it+1}(\ell', r')^{(h)}; \theta_u^{(h+1)}, \hat{\theta}_g, \hat{\sigma}_g(a), \hat{\theta}_r, \hat{\sigma}_r \right) \right)} \tag{21}$$

We continue these two steps until convergence.

7.3 Estimation With Unobserved Heterogeneity

After accounting for unobserved heterogeneity, we introduce additional parameters: α_k, β_k, and θ_k. These parameters capture the unobserved heterogeneity entering the utility, the unobserved heterogeneity entering the GPA production, and the probability of being unobserved type k conditional on the initial condition x_{i1}, respectively. At the start of iteration $(h+1)$, we first update the posterior ability distributions using all observed outcomes and course choice data, the production and learning parameters, and the Bayesian updating formulas for the posterior ability mean and covariance. With the obtained pdf of the posterior ability distribution, we construct the type-specific log-likelihood associated with the transition processes and the utility. We then calculate the posterior probability of being type k as follows:

$$q_i(k)^{(h)} = \frac{\pi^{(h)}_{k|x_{i1}} \exp\left[\sum_{t,a}(a_{it}=a)\mathbb{E}(L_{g,it}(a,k)^{(h)})\right] \exp\left[\sum_{t,a}(a_{it}=a)L_{u,it}(a,k)^{(h)}\right]}{\sum_k \pi^{(h)}_{k|x_{i1}} \exp\left[\sum_{t,a}(a_{it}=a)\mathbb{E}(L_{g,it}(a,k)^{(h)})\right] \exp\left[\sum_{t,a}(a_{it}=a)L_{u,it}(a,k)^{(h)}\right]} \tag{22}$$

where $\pi^{(h)}_{k|x_{i1}}$ denotes the probability of a student i being unobserved type k conditional on their initial conditions x_{i1}:

$$\pi^{(h)}_{k|x_{i1}} = \frac{\exp(x_{i1}\theta^{(h)}_k)}{\sum_k \exp(x_{i1}\theta^{(h)}_k)}. \tag{23}$$

We proceed to update the parameters governing the probability of student i being type k conditional on their initial conditions x_{i1}:

$$\theta^{(h+1)}_k = \arg max_{\theta_k} \sum_{i,k} q_i(k)^{(h)} \log\left(\pi_{k|x_{i1}}(\theta_k)\right) \tag{24}$$

Next, we update the population variance of the ability distribution:

$$(\sigma_g(a)^{(h+1)})^2 = \frac{1}{N}\sum_i \left((\sigma_{g,i}(a)^{(h)})^2 + \sum_k q_i(k)^{(h)}(\mu_{g,i}(a,k)^{(h)})^2\right) \tag{25}$$

where N is the sample size. We then update the parameters governing the GPA transition process:

$$\theta^{(h+1)}_g = \arg max_{\theta_g} \sum_{i,k} q_i(k)^{(h)} \sum_{t,a}(a_{it}=a)\mathbb{E}\left(L_{g,it}(a,k)^{(h)}\right) \tag{26}$$

followed by updating the parameters governing the utility:

$$\theta^{(h+1)}_u = \arg max_{\theta_u} \sum_{i,k} q_i(k)^{(h)} \sum_{t,a}(a_{it}=a)L_{u,it}(a,k)^{(h)} \tag{27}$$

Finally, we update the type-specific conditional choice probabilities. We repeat these steps until convergence is achieved. We summarize the estimation steps in 1.

8 Model Estimates

In this section, we present the estimated parameters of our dynamic discrete choice model of high school course-taking. We focus on the key components of the model, including the flow utility parameters, rank production and learning parameters, and GPA production and learning parameters. We also assess the model's performance in capturing observed choice patterns using a forward simulation approach.

Algorithm 1 Sequential E-M Algorithm

Require: Initial guesses for parameters and CCPs
Ensure: Convergence of parameters and CCPs
Initialize iteration counter $h \leftarrow 1$
while not converged **do**
Construct the posterior probability of being type k using Equation (22), denoted by $q_i(k)^{(h)}$
Update the parameters governing the probability of being type k $\theta_k^{(h+1)} \leftarrow \arg\max_{\theta_k} \sum_{i,k} q_i(k)^{(h)} \log\left(\pi_{k|x_{i1}}(\theta_k)\right)$
Update population variance: $(\sigma_g(a)^{(h)})^2 \leftarrow \frac{1}{N}\sum_{i=1}^{N}(\sigma_{g,i}(a)^{(h)})^2 + \sum_k q_i(k)^{(h)}(\mu_{g,i}(a,k)^{(h)})^2$
Update the GPA production parameters $\theta_g^{(h+1)} \leftarrow \arg\max_{\theta_g} \sum_{i,k} q_i(k)^{(h)} \sum_{t,a}(a_{it}=a)\,\mathbb{E}\left(L_{g,it}(a,k)^{(h)}\right)$
Update the flow utility parameters $\theta_u^{(h+1)} \leftarrow \arg\max_{\theta_u} \sum q_i(k)^{(h)} \sum (a_{it}=a)\, L_{u,it}(a,k)^{(h)}$
Update the CCPs $p_{it}(a,\ell,r,k)^{(h+1)} \leftarrow \frac{\exp\left(v_t\left(a,\ell,r,k,p_{it+1}(\ell',r',k)^{(h)};\theta_u^{(h+1)},\theta_g^{(h+1)},\sigma_g(a)^{(h+1)},\hat{\theta}_r,\hat{\sigma}_r\right)\right)}{\sum_a \exp\left(v_t\left(a,\ell,r,k,p_{it+1}(\ell',r',k)^{(h)};\theta_u^{(h+1)},\theta_g^{(h+1)},\sigma_g(a)^{(h+1)},\hat{\theta}_r,\hat{\sigma}_r\right)\right)}$
$h \leftarrow h+1$
end while=0

Table 6. Flow Utility Parameters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	None		Regular		Weighted		Accelerated	
	Est.	(S.E.)	Est.	(S.E.)	Est.	(S.E.)	Est.	(S.E.)
Constant	-2.86	(0.03)	-0.96	(0.03)	-7.15	(0.05)	-2.55	(0.03)
Accelerated	1.16	(0.02)	-1.33	(0.02)	0.77	(0.02)	-0.21	(0.03)
Past Rank	3.49	(0.10)	-1.16	(0.09)	0.25	(0.10)	-3.44	(0.11)
Past Rank squared	-2.81	(0.11)	1.58	(0.10)	1.09	(0.11)	4.35	(0.12)
Black	-0.16	(0.02)	-0.00	(0.01)	0.08	(0.02)	-0.17	(0.02)
Hispanic	-0.24	(0.03)	0.16	(0.02)	0.06	(0.03)	-0.37	(0.03)
Female	-0.13	(0.02)	0.27	(0.02)	0.26	(0.02)	-0.17	(0.03)
Type 2	-0.85	(0.02)	0.13	(0.02)	5.16	(0.05)	0.25	(0.02)

*Note:*This table presents estimated coefficients and standard errors for various student characteristics' effects on course placement across different alternatives (None, Regular, Weighted, Accelerated). The dependent variable indicates which alternative is chosen. Coefficients are reported for variables including accelerated status, past rank, race/ethnicity, gender, and student type. Standard errors are shown in parentheses.
Data source: Administrative data from a large urban district in Texas

8.1 Flow Utility Parameters

Table 1 reports the estimated flow utility parameters for each course option: None, Regular, Weighted, and Accelerated. The constant term is highest for the Regular course (-0.96) and lowest for the Weighted course (-7.15), suggesting that weighted courses may involve a higher cost of effort for students. The Accelerated dummy variable, indicating whether a student is above the recommended math level for their grade, has a positive effect on the utility of choosing Weighted courses (0.77) but a negative effect on Regular (-1.33) and Accelerated (-0.21) courses. This suggests that accelerated students may prefer weighted courses to maintain their advanced standing while avoiding the additional challenge of accelerated courses or the repetition of regular courses.

Demographic factors also play a role in shaping students' course preferences. Being Black or Hispanic is negatively associated with the utility of choosing Accelerated courses, while being female is positively associated with the utility of choosing Regular and Weighted courses. The model accounts for unobserved student heterogeneity by incorporating two types, with Type 2 students exhibiting a significantly higher utility for Weighted courses (5.16) compared to other options, indicating systematic differences in motivation and preferences.

8.2 Rank Production and Learning Parameters

The rank production function parameter estimates reveal that a one standard deviation increase in GPA leads to a 1.78-unit increase in rank, underscoring the strong positive association between individual achievement and relative position within the cohort. The standard deviation of the noise in the rank production function is estimated to be 0.34, with a signal-to-noise ratio of 0.097, suggesting that rank signals are relatively noisy and may not provide highly precise information about a student's true relative math ability.

8.3 GPA Production and Learning Parameters

Table 2 presents the estimated GPA production function parameters for each course option. Having taken an accelerated course in the previous period has a positive effect on GPA for students who subsequently enroll in a Weighted course (0.13) but a slightly negative effect for those who take a Regular course (-0.04). Past GPA emerges as a strong predictor of current GPA across all course options, with coefficients ranging from 0.75 for Accelerated to 0.83 for Weighted courses.

The estimates also reveal notable patterns across demographic groups and unobserved student types. Hispanic and female students exhibit lower GPAs in Weighted and Accelerated courses compared to their counterparts. Type 2 students, on average, have a higher GPA across all course options, aligning with their higher utility from taking Weighted courses and suggesting that they possess unobserved characteristics contributing to superior academic performance.

The standard deviation of the noise in the GPA production function varies across course options, with the Accelerated course exhibiting the highest noise standard deviation (0.79). The signal-to-noise ratio, measuring the informativeness of GPA signals, is highest for the Accelerated course

Table 7. GPA Production and Learning Parameters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	None		Regular		Weighted		Accelerated	
	Est.	(S.E.)	Est.	(S.E.)	Est.	(S.E.)	Est.	(S.E.)
Constant	-0.10	(0.01)	-0.14	(0.00)	-0.11	(0.02)	-0.01	(0.01)
Accelerated	-0.19	(0.00)	-0.04	(0.01)	0.13	(0.00)	0.01	(0.01)
Past GPA	0.81	(0.00)	0.78	(0.00)	0.83	(0.00)	0.75	(0.01)
Black	0.11	(0.00)	0.14	(0.00)	0.10	(0.00)	0.12	(0.01)
Hispanic	-0.01	(0.01)	0.00	(0.01)	-0.09	(0.01)	-0.10	(0.02)
Female	0.02	(0.01)	0.01	(0.00)	-0.05	(0.01)	-0.09	(0.01)
Type 2	0.29	(0.00)	0.29	(0.00)	0.05	(0.02)	0.02	(0.01)
Noise S.D.	0.43	(0.00)	0.66	(0.00)	0.57	(0.00)	0.79	(0.01)
Signal-to-Noise Ratio	0.05		0.02		0.10		0.23	

*Note:*This table presents estimated coefficients and standard errors for various student characteristics' effects on future GPA across different alternatives (None, Regular, Weighted, Accelerated). The dependent variable indicates future GPA. Coefficients are reported for variables including accelerated status, past GPA, race/ethnicity, gender, and student type. Standard errors are shown in parentheses.
Data source: Administrative data from a large urban district in Texas

(0.23) and lowest for the Regular course (0.02).

8.4 Model Fit

Table 8

	(1)	(2)	(3)
	Grade 10	Grade 11	Grade 12
Leave	0.00	-0.04	-0.02
None	0.10	-0.01	-0.07
Regular	-0.05	0.07	0.07
Weighted	-0.02	-0.02	0.02
Accelerated	-0.03	0.00	0.00

Note: Model frequencies are constructed using 10 simulations of the structural model for each individual included in the estimation.

To assess the model's performance, we employ a forward simulation approach, generating model-predicted choice probabilities and comparing them to actual choice probabilities in the data. The simulation process involves drawing individual and relative ability vectors, assigning unobserved types, generating preference shocks, and updating the state space based on the estimated structural parameters.

Table x presents a comparison of the model-predicted and actual choice probabilities by grade level and choice alternative, demonstrating that the model captures the choice patterns in the data remarkably well across all grades and alternatives.

9 Implications of Rank Non-Disclosure

In this section, we use the structural parameter and learning estimates to investigate the importance of information about one's abilities in the counterfactual scenario where students are not informed about their rank.

9.1 Counterfactual Policy Simulations

Table 9. Course enrollment frequencies in baseline and counterfactual: heterogeneity by clas rank

	(1)	(2)	(3)
	Grade 10	Grade 11	Grade 12
Panel A: Below 25th percentile			
Leave	-0.33	0.03	0.04
None	0.12	0.04	0.04
Regular	0.03	-0.02	-0.02
Weighted	0.06	-0.04	-0.12
Accelerated	0.02	0.03	0.05
Panel B: 25th to 50th percentile			
Leave	-0.33	0.02	0.02
None	0.08	0.02	0.03
Regular	0.02	-0.01	0.00
Weighted	0.04	-0.03	-0.05
Accelerated	0.03	0.01	0.01
Panel C: 50th to 75th percentile			
Leave	-0.37	-0.01	-0.01
None	0.04	0.00	-0.02
Regular	0.02	0.00	0.00
Weighted	0.04	0.01	0.03
Accelerated	0.03	0.01	0.02
Panel D: Above 75th percentile			
Leave	-0.44	-0.02	-0.04
None	0.05	-0.02	-0.09
Regular	0.00	-0.02	-0.02
Weighted	0.04	0.03	0.07
Accelerated	-0.01	0.03	0.04

Note: Model frequencies are constructed using 10 simulations of the structural model for each individual included in the estimation. Counterfactual frequencies use 10 simulations of each counterfactual model. "No disclosure" refers to our counterfactual where individuals have no information about their class rank.

To simulate the counterfactual scenario, we consider three main sources of uncertainty: individual preference shocks, idiosyncratic shocks in the GPA and rank functions, and abilities. In addition, students are assumed to be unaware of their rank realization, although they continue to know their GPA realization. The absence of rank information has two implications for the model: (1) the student's perceived rank enters the flow utility, and (2) without rank signals, there is no further updating of beliefs about cohort abilities, so the belief about future rank is based on the student's initial prior.

We solve the model backwards to obtain the counterfactual choice probabilities and then forward simulate to obtain the distribution of choices and the average abilities across different choice paths. The results of these simulations are presented in Table x.

Table x displays the changes in choice probabilities under the rank non-disclosure policy, relative to the baseline scenario where students receive information about their rank. The results are disaggregated by grade level (10, 11, and 12) and by the student's initial rank percentile (below 25th, 25th to 50th, 50th to 75th, and above 75th).

Several key patterns emerge from the counterfactual simulations. First, across all rank percentiles, the probability of leaving school decreases substantially in grade 10 under the rank non-disclosure policy. This suggests that withholding rank information may encourage students to stay in school, particularly in the early years of high school.

Second, the impact of rank non-disclosure on course choice probabilities varies by initial rank percentile. For students below the 50th percentile (Panels A and B), the policy leads to a decrease in the probability of choosing weighted courses, particularly in grades 11 and 12. In contrast, for students above the 50th percentile (Panels C and D), the policy increases the probability of choosing weighted courses, especially in grade 12.

Third, the counterfactual simulations reveal heterogeneous effects of rank non-disclosure on the probability of choosing accelerated courses. While the policy increases the likelihood of choosing accelerated courses for students below the 50th percentile, it has mixed effects for students above the 50th percentile, with a decrease in grade 10 and an increase in grades 11 and 12.

These findings suggest that the impact of withholding rank information on students' course choices depends on their initial rank percentile. Students with lower initial ranks may be discouraged from taking more challenging courses in the absence of rank signals, while students with higher initial ranks may be more inclined to enroll in weighted and accelerated courses.

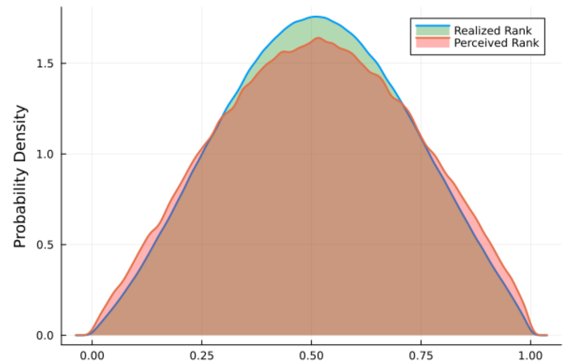
The counterfactual policy simulations provide valuable insights into the potential consequences of rank non-disclosure on students' educational decisions. The results highlight the importance of considering heterogeneous effects across different student subgroups when evaluating the impact of information provision policies.

9.2 Mechanisms

To better understand the ability sorting patterns, we examine the posterior means of the unobserved individual and cohort abilities in the period of last high school enrollment. These results, presented in Table x, are obtained using the same forward simulation process described in the previous section.

Two key patterns emerge from Table x. First, students who choose weighted courses have generally received strongly positive signals regarding their individual abilities. Second, these students

Figure 4. Distribution of Realized and Perceived Rank



Note: This figure displays the probability density distributions of realized rank (green) and perceived rank (red). The x-axis represents the rank from 0 to 1, where 0 is the lowest rank and 1 is the highest. The y-axis shows the probability density. The distributions largely overlap, with the perceived rank distribution showing a slight upward bias in the right-tail compared to the realized rank distribution. This suggests that individuals tend to perceive their rank as somewhat higher than their actual position, particularly in the right tail.

have also received strongly positive signals about their cohort's abilities.

Panel A of Table x shows the posterior means of individual ability for each course choice in grades 11 and 12. Students who choose weighted courses have the highest posterior means of individual ability in both grades (0.006 in grade 11 and 0.015 in grade 12), indicating that they have received positive signals about their own academic abilities.

Panel B of Table x presents the posterior means of cohort ability for each course choice. Again, students who choose weighted courses have the highest posterior means of cohort ability (0.007 in grade 11 and 0.017 in grade 12), suggesting that they have also received positive signals about the abilities of their peers.

These findings suggest that when students are not provided with information about their rank, which would have revealed their relative standing within the cohort, they are more likely to enroll in weighted courses if the withheld information would have led them to hold lower beliefs about their cohort's abilities. In other words, students who perceive their cohort as highly capable are more inclined to challenge themselves by taking weighted courses, even in the absence of explicit information about their relative rank.

The ability sorting patterns revealed in Table x provide valuable insights into the mechanisms driving students' course choice decisions. The results highlight the importance of both individual and cohort ability beliefs in shaping students' academic pursuits, and they underscore the potential impact of information provision on these decisions.

10 Quasi-Experimental Evidence

To complement our structural estimates, we rely on quasi-experimental evidence on the effect of rank non-disclosure on high school course choices. The objective of this analysis is to provide additional evidence - independent of our model's assumptions-on how information non-disclosure

Table 10. Ability Sorting

	(1)	(2)
	Grade 11	Grade 12
Panel A: Individual Ability		
None	0.000	0.000
Regular	-0.002	-0.003
Weighted	0.006	0.015
Accelerated	0.002	-0.001
Panel B: Cohort Ability		
None	-0.001	-0.001
Regular	0.001	0.003
Weighted	0.007	0.017
Accelerated	0.003	0.005

Note: This table presents the effects of ability sorting across different course types for both individual students and cohorts in Grades 11 and 12. Panel A shows individual ability effects, while Panel B displays cohort ability effects. Values represent the magnitude of sorting effects for each course type (None, Regular, Weighted, and Accelerated) relative to a baseline. Positive values indicate higher ability sorting, while negative values suggest lower ability sorting. Weighted courses show the strongest positive effect on ability sorting, particularly in Grade 12.

impacts high school course choices. The introduction of a rank non-disclosure in a school district in Texas provides us an opportunity to observe course choices in schools with and without rank disclosure.

Description Our analysis utilizes a comprehensive dataset encompassing 57 schools across three independent school districts in Texas, spanning nine school years and including a total of 142,279 students. The dataset is particularly valuable due to a recent policy change that occurred in District A, which implemented rank non-disclosure starting in year y. This policy change creates a clear distinction between treatment and control groups: students in District A from year y onward no longer received information about their class rank, while students in the other districts and in District A prior to year y continued to receive rank information. The dataset provides rich information on student demographics, course selections, GPA, and rank for each school year.

Table X presents summary statistics for key variables across the three districts, revealing notable differences in student composition and academic outcomes.

Panel A of Table X highlights demographic variations among the districts. District A has a notably lower proportion of Black and Hispanic students (11% and 15% respectively) compared to Districts B (24% and 72%) and C (24% and 67%). This demographic composition suggests that District A serves a more predominantly white or Asian student population.

Panel B reveals distinct patterns in math course-taking behavior across the three districts. Students in District A, on average, take more math courses (3.80) compared to those in District B (3.15) and C (3.39). More strikingly, the average number of weighted math courses taken in District A (1.98) is substantially higher than in Districts B (0.58) and C (0.97).

The data also show a clear trend in advanced math course completion rates. While almost all students across all districts complete Algebra I and over 90% complete Geometry, the completion rates for higher-level math courses are consistently higher in District A. For instance, 87% of students in District A complete Algebra II, compared to 73% and 80% in Districts B and C respectively. The gap widens further for Pre-Calculus (63% in A vs. 27% in B and 49% in C) and Calculus (29% in A vs. 4% in B and 14% in C).

Panel C presents dropout rates by grade level, revealing lower attrition in District A across all grades. By grade 10, only 6% of students in District A have dropped out, compared to 9% and 11% in Districts B and C. This trend persists through grades 11 and 12, with District A maintaining lower dropout rates. By grade 12, the cumulative dropout rate in District A (14%) is less than half that of District B (33%) and substantially lower than District C (29%).

Panel D illustrates consistently higher average GPAs in District A across all grade levels. This trend aligns with the observations from previous panels, further supporting the notion that District A students generally perform better academically.

It's important to note that District A is considerably smaller (8,769 students) compared to Districts B (37,750 students) and C (95,760 students). The combination of smaller size, higher course-taking rates in advanced math, lower dropout rates, and higher GPAs indicates that District A is, on average, a higher-performing district.

While these differences are noteworthy and could potentially confound our analysis, we account for these factors by including them as controls in our difference-in-differences approach to estimate the average treatment effect of the rank non-disclosure policy.

Analysis Our study investigates the impact of a rank non-disclosure policy implemented in District A in year y^*. We employ a difference-in-differences approach to estimate the average treatment effect on the treated (ATT) of this policy on students' likelihood of enrolling in weighted math courses. The treatment group is defined based on two key conditions: enrollment in District A and placement below the 90th percentile. This definition accounts for the state mandate requiring disclosure of rank information to students in the top 10th percentile, irrespective of local policies.

To capture the timing of policy exposure, we introduce the variable $Post_{itsy} = y^* - (t - 9)$, indicating whether a student entered grade level t after the policy implementation. Our analysis utilizes the difference-in-differences approach developed by Callaway and Sant'Anna (2021), which effectively accounts for the staggered nature of treatment exposure across academic years. This method is particularly suitable for our context, where students were exposed to the rank non-disclosure policy at different stages of their academic journey.

To enhance the comparability of our treatment and control groups, we impose two key sample restrictions. First, we limit the sample to students at or above the 75th percentile. This allows us to compare students who are similar in terms of academic performance, with those below the 90th percentile in the treatment group and those above in the control group. Second, we restrict the sample to students who have completed at most the recommended level of math. This addresses

Table 11. Past academic and background characteristics by course choice

	(1)	(2)	(3)
	District A	District B	District C
Panel A: Demographics			
Black	0.11	0.24	0.24
Hispanic	0.15	0.72	0.67
Female	0.49	0.49	0.50
Panel B: Math Coursetaking			
Total Math courses	3.80	3.15	3.39
Total Weighted Math courses	1.98	0.58	0.97
Ever Completed Algebra I	1.00	0.99	1.00
Ever Completed Geometry	0.95	0.92	0.92
Ever Completed Algebra II	0.87	0.73	0.80
Ever Completed Pre-Calculus	0.63	0.27	0.49
Ever Completed Calculus	0.29	0.04	0.14
Panel C: Dropout			
Dropout by Grade 9	-	-	-
Dropout by Grade 10	0.06	0.09	0.11
Dropout by Grade 11	0.11	0.18	0.21
Dropout by Grade 12	0.14	0.33	0.29
Panel D: GPA			
GPA after Grade 9	91.07	87.47	85.07
GPA after Grade 10	91.29	87.58	85.33
GPA after Grade 11	91.24	87.80	85.67
GPA after Grade 12	91.16	87.90	86.24
Obs	8,769	37,750	95,760

Note: This table reports summary statistics for the data that is used to estimate our structural model. This is student-level data.
Data source: Administrative data from a large urban district in Texas

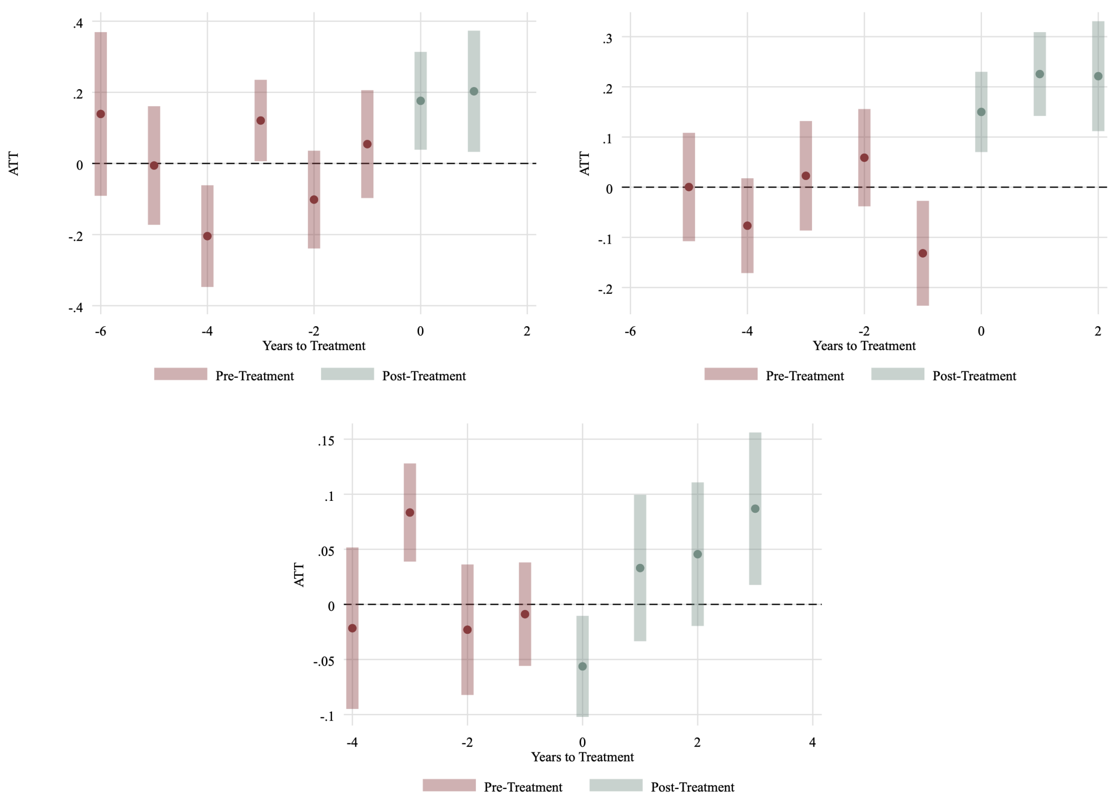
the systematic differences observed in District A, which appears to be higher-performing overall. These restrictions ensure that we are comparing similar high-performing students and controlling for District A's apparent higher overall performance. In our analysis, we also control for student demographics and cluster standard errors at the school level.

Figure X presents event study plots for grades 10, 11, and 12, illustrating the policy's impact on weighted math course enrollment. For grade 10 (Figure X(a)), we observe two post-treatment years, with the second cohort receiving full treatment (exposed from ninth grade). The results show a significant positive ATT in both post-treatment years, with no substantial difference between ninth and tenth-grade exposure. This similarity is likely because rank information is first received at the start of tenth grade, so exposure to the policy in ninth grade does not significantly alter the student's information set.

The grade 11 analysis (Figure X(b)) displays three post-treatment years, revealing a positive and significant ATT for all cohorts exposed to the policy. For grade 12 (Figure X(c)), we see four post-treatment years, with a significant positive ATT only for the fourth cohort that received full treatment starting from ninth grade. Interestingly, the ATT is close to zero or even negative for cohorts with shorter exposure periods.

Across all grades, we observe that pre-treatment trends are generally stable and close to zero, supporting the parallel trends assumption crucial for the validity of our difference-in-differences approach. However, we note a notable exception of a positive trend three years pre-implementation

Figure 5. Average Treatment Effects Over Time for Three Different Cohorts



Note: These graphs show the average treatment effects (ATT) over time for three different cohorts. The x-axis represents years relative to treatment, with 0 being the year of treatment. The y-axis shows the magnitude of the treatment effect. Red bars and points represent pre-treatment periods, while green bars and points represent post-treatment periods. The vertical bars indicate 95% confidence intervals.
Data: Administrative data from Independent School Districts in Texas

for grades 10 and 12, which warrants further investigation.

Table X summarizes the ATT for full treatment cohorts across grades, revealing substantial effects. Tenth graders are 20 percentage points more likely to take weighted math when they do not know their rank compared to when they do. The effect is even larger for eleventh graders, with a 22 percentage point increase, while twelfth graders show a 9 percentage point increase. These results suggest that the rank non-disclosure policy has a substantial positive impact on students' propensity to enroll in weighted math courses, with the effect varying by grade level and duration of exposure. The diminishing effect in grade 12 may indicate a ceiling effect or a shift in student priorities as they approach graduation.

These results are consistent with the model predicts and suggest that removing potential discouragement effects associated with rank disclosure may encourage students to pursue more rigorous coursework.

Table 12. Average Treatment Effect on Weighed Math Enrollment

	(1) Grade 10	(2) Grade 11	(3) Grade 12
ATT	0.20^{**} (0.09)	0.22^{***} (0.06)	0.09^{**} (0.04)
Mean	0.40	0.30	0.17
Obs	35,717	46,455	56,541

Note: This table reports the Average Treatment Effect (ATT) on weight math enrollment for grades 10, 11, and 12. Standard errors are reported in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$. Mean represents the average enrollment rate for each grade.
Data source: Administrative data from Independent School Districts in Texas

11 Conclusion

References

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1-2), 343–375.

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2025). College attrition and the dynamics of information revelation. *Journal of Political Economy*.

Arcidiacono, P., E. M. Aucejo, and K. Spenner (2012). What happens after enrollment? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics 1*, 1–24.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Arcidiacono, P. and R. A. Miller (2019). Nonstationary dynamic models with finite dependence. *Quantitative Economics 10*(3), 853–890.

Aughinbaugh, A. (2012). The effects of high school math curriculum on college attendance: Evidence from the nlsy97. *Economics of Education Review 31*(6), 861–870.

Azmat, G. and N. Iriberri (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics 94*(7-8), 435–452.

Balingit, M. (2015). High schools are doing away with class rank. what does that mean for college admissions?

Bertoni, M. and R. Nisticò (2023). Ordinal rank and the structure of ability peer effects. *Journal of Public Economics 217*, 104797.

Black, S. E., J. T. Denning, and J. Rothstein (2020). Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. Technical report, National Bureau of Economic Research.

Brade, R., O. Himmler, and R. Jäckle (2022). Relative performance feedback and the effects of being above average—field experiment and replication. *Economics of Education Review 89*, 102268.

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

Cicala, S., R. G. Fryer, J. L. Spenkuch, et al. (2011). A roy model of social interactions. Technical report, National Bureau of Economic Research.

Cortes, K. E., A. I. Friedson, et al. (2014). Ranking up by moving out: the effect of the texas top 10% plan on property values. *National Tax Journal 67*(1), 51–76.

Cullen, J. B., M. C. Long, and R. Reback (2013). Jockeying for position: Strategic high school choice under texas' top ten percent plan. *Journal of Public Economics 97*, 32–48.

Daugherty, L., P. Martorell, and I. McFarlin (2014). Percent plans, automatic admissions, and college outcomes. *IZA Journal of Labor Economics 3*(1), 1–29.

De Groote, O. (2019). A dynamic model of effort choice in high school.

Delaney, J. and P. J. Devereux (2022). Rank effects in education: What do we know so far?

Dellinger, H. (2022). Clear creek latest texas school district to change class rank system, dropping electives penalties. *Houston Chronicle*.

Denning, J. T., R. Murphy, and F. Weinhardt (2023). Class rank and long-run outcomes. *Review of Economics and Statistics 105*(6), 1426–1441.

Donaldson, E. (2020). North east isd to end class rankings for students not in top 10%. *San Antonio Report*.

Eckstein, Z. and K. I. Wolpin (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica 67*(6), 1295–1339.

Elsner, B. and I. E. Isphording (2017). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics 35*(3), 787–828.

Erdem, T. and M. P. Keane (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science 15*(1), 1–20.

Floyd, E., S. Tomar, and D. J. Lee (2024). Making the grade (but not disclosing it): How withholding grades affects student behavior and employment. *Management Science 70*(4), 2497–2517.

Genakos, C. and M. Pagliero (2012). Interim rank, risk taking, and performance in dynamic tournaments. *Journal of Political Economy 120*(4), 782–813.

Goulas, S. and R. Megalokonomou (2021). Knowing who you actually are: The effect of feedback on short-and longer-term outcomes. *Journal of Economic Behavior & Organization 183*, 589–615.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Hu, Y. and M. Shum (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics 171*(1), 32–44.

Hui, T. K. (2022). Why some nc families say 'class rank' is harmful and want it taken off school transcripts. *The News & Observer (Raleigh, N.C.)*.

James, J. (2011). Ability matching and occupational choice.

Jiang, Z. Z., Y. Huang, and D. R. Beil (2018). The role of feedback in dynamic crowdsourcing contests: A structural empirical analysis. *Ross School of Business Paper* (1334).

Kasahara, H. and K. Shimotsu (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica 77*(1), 135–175.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of political Economy 105*(3), 473–522.

Lemus, J. and G. Marshall (2021). Dynamic tournament design: Evidence from prediction contests. *Journal of Political Economy 129*(2), 000–000.

Miller, R. A. (1984). Job matching and occupational choice. *Journal of Political economy 92*(6), 1086–1120.

Murphy, R. and F. Weinhardt (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies 87*(6), 2777–2826.

Niu, S. X. and M. Tienda (2010). The impact of the texas top ten percent law on college enrollment: A regression discontinuity approach. *Journal of Policy Analysis and Management 29*(1), 84–110.

Office of House Bill Analysis (2001). *Senate Bill 974 HBA-NRS S.B. 974 77(R)*.

Potter, P. (2018). Unhealthy competition among reasons for changing high school class ranking in frisco isd. *The Dallas Morning News*.

Rose, H. and J. R. Betts (2004). The effect of high school courses on earnings. *Review of economics and statistics 86*(2), 497–513.

Rozen, S. (2022). Friendswood isd sees jump in college admissions with elimination of class rank. *Community Impact*.

Tchuente, G. (2016). High school human capital portfolio and college outcomes. *Journal of Human Capital 10*(3), 267–302.

The Texas Senate (2001). College admission bill creates a stir. *Texas Senate News*.

Todd, P. and S. Wang (2021). Understanding gender disparities in stem major and occupation choices: A random forest approach. *University of Pennsylvania*.

Tran, A. and R. Zeckhauser (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics 96*(9-10), 645–650.

Webb, S. (2019a). Frustrated with top 10% rule, houston-area schools stop ranking some students. *Houston Chronicle*.

Webb, S. (2019b). Katy isd eliminates d grades, changes how it reports class rankings. *Houston Chronicle*.

Zax, J. S. and D. I. Rees (2002). Iq, academic performance, environment, and earnings. *Review of Economics and Statistics 84*(4), 600–616.