Matching Students and Instructors in Higher Ed.*

León F. Guzmán Lizardo†

November 6, 2023

Please click here for the latest version.

Abstract

I study the measurement of matching effects in higher-education learning technologies, with an emphasis on understanding the learning implications of common course-enrollment mechanisms used to assign students to professors within a course. To achieve this, I construct an empirical model that describes a student's academic path in a standard higher-education institution. Two main conceptual contributions emerge. First, I show how to use sequences of subject-related courses to identify learning production functions when instructors differ in their grading policies, a situation common to most post-secondary settings. Second, I propose a new channel through which heterogeneity in instructors' grading policies can indirectly impact learning outcomes by influencing students' demands for instructors in choice-based course-enrollment mechanisms. To illustrate the content behind these results, I estimate the model using academic records from INTEC, a university in the Dominican Republic. The estimates reveal substantial student-professor matching effects and a strong student preference for expected scores, relative to expected learning, under a potential professor match. By constructing a counterfactual exercise that reassigns students and professors within a course, I demonstrate how the assignment resulting from INTEC's course-enrollment mechanism, based on both random and choice-based rules, can be significantly improved in terms of learning output, course-dropout rates, and the number of course retakes.

*I am grateful to my advisors Alfred Galichon, Quang Vuong, and Daniel Waldinger for their support and guidance.

†Department of Economics, New York University. Email: lgl266@nyu.edu.

1 Introduction

Educational institutions are in charge of pairing students to professors so that teaching can take place. If a learning concern is behind this task, understanding the extent to which learning technologies are subject to matching effects is key to the construction of these matches. For example, knowing which instructors are more effective at facilitating learning for disadvantaged students, relative to advanced ones, is a useful input. Despite its significance, the latter has received limited attention in the empirical literature, with the few existing studies primarily focusing on settings other than post-secondary environments. This distinction is relevant not only because of the substantial amount of learning that occurs in post-secondary settings but also due to the distinct structural features describing them. In particular, the fact that student-professor assignments are typically constructed through course-enrollment mechanisms which, while serving other institutional goals, have learning consequences that are not well understood.

This paper examines the measurement of student-professor matching effects in higher-education settings and studies how commonly used course-enrollment mechanisms can result in inefficient learning matches. To achieve this, I construct and estimate a structural model describing the learning outcomes a student achieves throughout a sequence of courses at a typical post-secondary institution. In addition to describing the nature of the observed outcomes, the estimates are used in the evaluation of counterfactual policies aiming to enhance learning outcomes through reassignments. Two main conceptual contributions result from this exercise.

First, I provide a collection of arguments for the identification of learning technologies in situations where instructors differ in the grading policies they use to evaluate students' learning. The arguments allow for a very rich class of learning technologies, with the only requirement imposed on learning production functions being injectivity with respect to its inputs. The primary technical challenge arises from the need to disentangle an instructor's teaching effectiveness and grading policy from the observed distribution of scores. The results show how to address this issue by using sequences of subject-related mandatory courses, a common feature in most higher-education institutions.

To illustrate these ideas, consider the task of comparing the teaching quality of two Calculus 1 professors. In the absence of grading policy differences across instructors, a straightforward thought experiment reveals how to proceed: (i) select two students with identical histories, (ii) assign each to one of the two professors under consideration, and (iii) compare the instructors in terms of the differences in the students' Calculus 1 scores. Indeed this approach, based on within-professor variations of the observed scores, is followed by a substantial portion of the literature focusing on elementary and secondary education settings where standardized tests are readily available. However, the approach becomes problematic in situations where instructors, in addition to their teaching abilities, differ in their grading policies. This is because variations in the observed scores across instructors could arise from either differences in their teaching abilities or disparities in how they choose to grade the course. Hence, it becomes obvious that the comparison should rely on an outcome related to learning in Calculus 1 but independent of Calculus 1's grading policies.

The structure of the post-secondary curriculum offers a logical alternative. Specifically, we can employ sequences of subject-related mandatory courses within a major program to evaluate a professor's teaching impact in a given course in terms of the performance of its students in the subsequent courses within the sequence. Intuitively, while a student's performance in later courses of the sequence depends on its learning in the preceding courses, it remains unaffected by the grading policies of the latter. In terms of our thought experiment, we can consider assigning both students to a common Calculus 2 instructor and to use the difference in their score performance to infer the teaching ability differences among the Calculus 1 instructors.

The second contribution explores the learning properties of the student-to-professor assignments resulting from different course-enrollment mechanisms. In particular, I study the properties of (i) random assignment and (ii) first-come-first-serve mechanisms, both commonly used in many post-secondary institutions including my empirical setting. For the second, I propose a new channel through which the nature of grading policies can affect learning outcomes indirectly by modifying the assignment of students to professors. Intuitively, given the emphasis on assessing student performance through grades and the heterogeneity in grading policies among instructors, the way a student participates in a choice-based course-enrollment mechanism is likely influenced not only by a concern for learning, but also by a preference for achieving high scores. If high-learning instructors do not coincide with high-scoring professors, these mechanisms are unlikely to result in learning-efficient assignments.

To illustrate the content of these results, I estimate the model using the academic records of INTEC, a post-secondary education institution in the Dominican Republic. The empirical exercises focus on the initial two courses of the Calculus sequence offered by INTEC (i.e., Calculus 1 and 2). Two key insights can be derived from the model's estimates. First, while students' inputs play an important role in predicting learning outcomes, these heavily depend on the student's professor match. For example, among students at the top of the ability distribution, as measured by their scores in the math component of INTEC's entrance exam, the learning gap between the best and worst professors for a given student can correspond to more than a full letter score on the GPA scale. Second, there is substantial variation in the grading policies employed by instructors, particularly in terms of the marginal score return associated with learning. Notably, instructors with lenient grading policies do not always coincide with high-learning instructors. This implies a score vs. learning tension for students demanding sections within a course.

The latter tension can be captured through a straightforward exercise in which I construct estimates for the learning/scoring gaps resulting from a match with the learning-optimal and scoring-optimal instructors for each course-enrollment instance observed in the sample. For instance, by enrolling the associated scoring-optimal instructor as opposed to the learning-optimal instructor, the average student faces a learning opportunity cost of 0.17 GPA points. If instead the student were to enroll the learning-optimal instructor's section the student would face a scoring opportunity cost of 0.76 GPA points relative to what it would obtain under the scoring-optimal instructor. I argue that this is a large difference with important consequences over how a student chooses to demand a section within a course when allowed to do so. In particular given that the estimates suggest students place a higher weight on scoring outcomes, relative to learning outcomes, when choosing among available sections. Indeed, among the students who can select their professors by participating in the first-come-first-serve mechanism, approximately 83.2% of them end up matched with an instructor other than the learning-optimal one for them according to the estimates.

Regarding the policy implications of these findings, I explore a collection of counterfactual policies involving a dean seeking to assign students to professors in a dictatorial manner. Various versions of this exercise are considered, each assigning different weights to students based on their initial abilities as a way of capturing different distributional concerns in designing a student-professor match. The policy simulations reveal significant gains under all choices of the dean's weighting function. These gains are evident both directly in terms of our learning definition and in terms of other learning-related variables, such as student dropout rates and the number of retakes required to achieve a passing score. Importantly, while the gains vary in magnitude, they are positive across the entire distribution of student ability, suggesting that the university can improve efficiency in the production of learning without sacrificing distributional concerns.

The remainder of this paper is organized as follows. Section 2 starts by locating the contributions of the paper within the education literature. This is followed by Section 3 which introduces the empirical setting considered in this paper and describes some stylized facts key to the construction of the empirical model. Section 5 introduces the conceptual model, dividing the discussion in terms of a model for the learning production function and a model for the demand of sections within a course. Section 6, introduces the empirical model and the arguments behind the identification of the primitives of interest. Section 7 documents the estimates of the model and discusses some of the implications resulting from them. Finally, section 8 considers the results of the counterfactual reassignment policies. Section 9 concludes and suggests some future research avenues.

2 Related Literature

This project contributes to an extensive body of research studying heterogeneity in instructors' teaching quality. Within this group a significant amount of work has taken place under the value-added framework. Some early contributions include Hanushek (1971), Rockoff (2004), Rivkin et al. (2005), and Hanushek (2009). These are primarily empirical projects emphasizing the use of teacher quality measures as a guide for hiring, promotion, and dismissal decisions in education settings. They document substantial disparities among instructors' teaching quality and gains from policies that act on such differences. In terms of recent conceptual and methodological contributions, consider Kane & Staiger (2008), Chetty et al. (2014),Gilraine et al. (2020), and Gilraine & Pope (2021).

Common to these papers is the assumption that learning production functions additively separate student and professor inputs. This modeling decision, stemming from a focus on describing differences in the average teaching quality across instructors, is reasonable given the policy questions these papers seek to answer. However, it is inadequate for the purpose of studying assignment problems since, under separability, aggregate learning outcomes are independent of the way in which instructors and students are matched. Recent value-added examples exploring the existence of matching effects in the learning technology can be found in Aucejo et al. (2018), Ahn et al. (2020), and Graham et al. (2022). Although these projects share the concern for matching effects in the learning technology, they differ from this paper in several important ways.

The most obvious is my focus on higher-education settings. This distinction is significant because the structural differences between higher-education settings and other learning environments rule out any simple extrapolation of the conclusions derived from estimates based on the latter. Instead, these call for different modeling and econometric approaches. Examples of these differences include the heterogeneity in the grading policies used by professors to map learning into scores, the fact that only discrete scores are reported (i.e., letter scores), and the truncation in the distribution of scores resulting from students being able to drop sections of a course. Relative to value-added papers that do consider matching effects, the approach proposed here allows for richer forms of complementarities beyond the multiplicative separability used in many of these contributions: the only requirement is for the professor-specific learning production functions to be injective relative to its inputs. This is, of course, important, as functional restrictions on the learning production functions can impact the policy conclusions resulting from the estimates.

More important are the methodological differences in the identification of the matching effects. In particular, I show how the within-professor variations in the score distribution, used in elementary and secondary settings to identify learning technologies, cannot be used in higher-education settings where instructors differ in their grading policies. Intuitively, high scores under an instructor can result from either high teaching quality or a lenient grading policy. I provide arguments for disentangling both based on the observation of students in multiple periods along a sequence of subject-related courses. The approach is broad in that it can be applied to a large class of post-secondary institutions and is structural, which facilitates the evaluation of counterfactual policies.

The spirit of the identification argument is similar to that in Carrell & West (2010), which considers grades in future courses as the normalization defining learning. However, their concern is not one of confounding grading policies and learning, as their setting involves standardized tests. Instead, it is about distinguishing between instructors who "teach to the test" and those who have a lasting learning impact on students. The reduced form approach followed by this contribution allows for testing the hypothesis of pedagogical differences across instructors in a post-secondary education institution but does not lend itself to counterfactual analysis. Moreover, the use of the empirical arguments requires a setting with standardized testing and random assignment of students to professors, an uncommon situation in most higher-education institutions.

A second branch of the education literature related to this project has focused on understanding the consequences of differences in the grading policies across instructors on students' decisions within the university. Some recent examples are Ahn et al. (2019) and Butcher et al. (2014), focusing on students' decisions regarding major choices and how these are impacted by grading policies. Another example, Babcock (2010), focuses on student effort decisions within the course as an optimal response to the way in which grading policies affect the marginal returns to studying for students who value scores as an output. Like these papers, I adopt the perspective of grading policies potentially impacting student decisions. However, the nature of the decisions considered here is very different. Specifically, I propose a new channel through which heterogeneity of grading policies can affect learning indirectly by modifying students' course/section demand decisions and, therefore, the resulting assignment. This must be contrasted with the emphasis on describing distortions resulting from students' choices of major, courses, and effort levels followed by these contributions.

Although related, these are conceptually very different. For instance, even in the absence of matching effects in the learning technology, major and course choice distortions might arise if grading policy differences exist across courses or departments. Conversely, distortions in the assignment of students to professors within a course can arise even in environments where students have little control over their major or course requirements (as is the case in my empirical setting after the initial enrollment). In a sense, the focus on matching effects, which requires to directly model the production technology, places this project at the intersection between the grading policy heterogeneity effects literature and the literature concerned with quantifying real learning differences across instructors mentioned above. Recent examples that also directly model the learning technology are Gershenson et al. (2022) and Figlio & Lucas (2004), documenting real learning consequences of grading policies adopted by professors. However, in these papers, the focus is again on elementary and secondary settings under technologies that don't factor in student-professor complementarities.

Finally, this project can be related to the literature studying assignment mechanisms, particularly those concerned with course-allocation problems. Although most of the papers here are of a theoretical nature, I share the concern of considering course-enrollment mechanisms that the university can directly choose to achieve different objectives. Some examples directly addressing the assignment problem in educational settings are Diebold et al. (2014), Krishna & Ünver (2008), and Sönmez & Ünver (2010).

One main difference with these papers is that for them, the focus is almost exclusively on the comparison of allocation mechanisms under preference based criteria. For example, very often the idea is to set up rules that lead to a student/course assignment satisfying notions of efficiency, fairness, and stability in terms of students' preferences. This approach, while reasonable, is not the only one. One can easily entertain ranking the assignment mechanisms in terms of the learning outputs they induce. Indeed, learning considerations are very often the stated goals guiding the decisions of universities. The approach pursued here also differs in that instead of considering the construction of mechanisms satisfying certain properties, it seeks to compare existing mechanisms by using the estimates for the primitives in the model. An exception to the latter is Budish & Cantillon (2012) that considers, under a preference-based approach, the comparison of course-enrollment mechanisms used in a concrete empirical setting.

3 Institutional details and data sources

The dataset corresponds to the academic records of the Instituto Tecnológico de Santo Domingo (INTEC), a higher-education institution in Dominican Republic primarily oriented towards undergraduate STEM majors. Information is arranged as longitudinal panel tracking the academic path of all students who enroll at INTEC between 2007 and 2022. A large set of information describing each course-enrollment instance for these students is available. This includes unique identifiers for each course-section, identifiers for the leading instructors of these sections, student's course dropping decisions, and the final score achieved by the student in the course (i.e., contingent upon not dropping the course). In addition, learning-related covariates are available at both the section and student levels.

On the student side, the main variable is its score on each component of the entrance exam evaluation. This test is designed to assess a student's understanding of the high-school material serving as prerequisite for the undergraduate curriculum. The latter variable serves as an initial ability measurement in the empirical analysis. Turning to variables associated to sections of a course, I construct several key attributes for each academic period previously identified in the literature as relevant for learning outcomes. These include the leading professor's teaching load in each academic term, the instructors course specific tenure (defined as the number of academic terms the instructor has been teaching the course in question), and the instructor's general tenure within the university (defined as the number of academic periods the professor has been active in the university). Moreover, for a large subset of sections, I have access to additional pertinent characteristics, such as the schedule of the section and the highest academic degree attained by the lead instructor.

As is common practice in many of Latin America's higher-education institutions, students at INTEC must choose a major program upon enrollment in the university. After making this decision, students are required to adhere to a predetermined curriculum comprising a collection of sequences of subject-related mandatory courses connected by prerequisite relationships. This system should be contrasted to North America's higher-education model where students often specify their major only after the completion of the associated credit/course requirements. The rigid nature of the curriculum in this setting facilitates the empirical exercise by allowing me to bypass the complexities of the combinatorial problem arising when students can choose the courses they want to enroll and the academic term in which to enroll them.

The empirical exercise focuses on a specific sequence of courses: Calculus 1 and Calculus 2. This choice offers several advantages from an empirical standpoint. First, the content covered in this sequence is highly standardized, suggesting that differences in the quality of instruction are of a vertical, rather than horizontal, nature. Second, this sequence is an integral part of the curriculum for nearly all university students, meaning I can observe a large number of students participating in these courses. Third, students start enrolling courses in the sequence on their first academic term in the university, ensuring there is no gap between the assessment of a student's ability (e.g., entrance exam scores) and its initial enrollment in the Calculus sequence. Additionally, since Calculus serves as a foundational course for most other major subjects pursued by the students in the sample, the courses in the sequence are typically enrolled in isolation (relative to other math oriented courses), reducing potential interference bias resulting from students learning the Calculus curriculum from instructors in related courses.

The university follows a scoring system based on the standard 4.0 GPA scale which maps into letter scores according to institutional cutoffs (i.e., A, B+, B, C+, C, D, F, and R (indicating a course section drop)). Importantly, scores are determined on an absolute basis, with each instructor independently choosing the grading policy mapping student learning into letter scores. In other words, there is no university mandate explicitly directing professors to achieve a specific score distribution within each section. This grading approach aligns with cultural and political considerations at the university, where it is presumed that an instructor's grading conveys significance beyond mere within-section student ranking.

Table 1: Descriptive statistics - Course enrollment level

	Calculus 1		**Calculus 2**	
	Avg.	Std. Dev.	Avg.	Std. Dev.
All students	2.80	1.17	2.75	1.07
Ability 0% - 25%	2.41	1.21	2.45	1.09
Ability 25% - 50%	2.65	1.21	2.55	1.07
Ability 50% - 75%	2.82	1.13	2.69	1.07
Ability 75% - 100%	3.22	0.99	3.08	0.99
Female	2.94	1.10	2.87	1.01
Male	2.70	1.22	2.66	1.11
Stem	2.81	1.19	2.74	1.09
Social sciences	2.77	1.16	2.59	1.09
Health sciences	2.78	1.16	2.86	1.00
High load	2.70	1.17	2.78	1.07
Low load	2.90	1.17	2.73	1.08
High course tenure	2.96	1.10	2.85	1.04
Low course tenure	2.72	1.20	2.68	1.09
High tenure	2.92	1.11	2.85	1.05
Low tenure	2.72	1.21	2.67	1.09

Notes: Statistics in the table correspond to the course-enrollment instance level. In the final panel, high/low are defined in terms of above/below average for each of the variables being considered.

Table 1 shows the distribution of scores for different subpopulations of interest in the data. All descriptive statistics in the table are given at the course-enrollment instance level. Overall, the average score for Calculus 1 stands at approximately 2.80 points in the 4.00 on the GPA scale, while that for Calculus 2 is approximately 2.75. In both instances, these averages map to a letter grade of B, based on the institutional letter score cutoffs. Scores are described by substantial variation around these average values. For example, the unconditional standard deviation for students in Calculus 1 exceeds 1.00 GPA point, implying a window around the mean score ranging from a letter grade of D to a score A.

Each of the subsequent panels in the table decomposes the distribution with respect to a conditioning variable of interest. For example, the second panel conditions on the ability level of students as measured by the student's exam in the math component of the entrance exam score. For both courses, students with a higher ability level are associated to higher average scores and less dispersion. In turn the third panel shows the distribution of scores conditional on biological sex. In line with findings from prior research, average score for both Calculus 1 and Calculus 2 are higher for female students compared to their male counterparts. Another key variable on the student side is the major choice made by each student. Intuitively, given the varying levels of mathematical emphasis across different majors, one would anticipate differences in the score distribution based on major selection. I categorize all observed majors into four primary groups in a way that closely aligns with the university's own classification: (1) STEM majors, (2) business and social sciences majors, (3) health majors, and (4) other majors. Small differences exist across these categories. For instance, when considering Calculus 1, the average score ranges from 2.77 for the second group to 2.81 for the first group. Differences in Calculus 2 are larger where the average score ranges fro 2.59 to 2.86 in the GPA scale. It's essential to bear in mind that while interpreting these averages, part of the observed differences reflect variations in course drop rates among students across different majors.

Besides showing the distribution of scores conditional on student subpopulations, the last panel of Table 1 shows how the distribution of scores in each course varies as a function of the teaching load and the tenure of the leading instructor for the section enrolled by the student. In the context of Calculus 1, instructors with a below average teaching load tend to exhibit higher average scores, whereas the opposite pattern emerges in the case of Calculus 2. In terms of instructor's course-specific and general tenure, for both Calculus courses, above average tenure (i.e., both course-specific and general) is associated with higher scores and a smaller score dispersion.

The course-enrollment mechanism defining the assignment of students to instructors in INTEC can be described by two main rules. First, first-term students (i.e., students in their initial academic term), are assigned to a section of Calculus 1 directly by the administration. The administration claims this assignment is done either randomly of in terms of variables which are orthogonal to the learning process of students upon being paired. Second, after the first academic term students enroll sections of a course by participating of a first-come-first-serve course-selection mechanism. For such students, at the beginning of the term an online platform opens and allows students to choose sections of the courses they seek to enroll. Each section has a capacity constraint, and a student's choice set comprises all courses whose capacity constraint is not binding at the moment of enrollment. While I observe data on each student's choice, I do not directly observe students' platform entry time or the specific choice set available to a student upon entering the platform.

Table 2: Descriptive statistics - Course/section level

	Calculus 1		**Calculus 2**	
	Avg.	Std. Dev.	Avg.	Std. Dev.
Section size	33.35	9.17	32.89	10.54
% at capacity	27.00	–	40.00	–
Mean score	2.70	0.69	2.66	0.56
Pass rate	0.63	0.24	0.63	0.22
Drop rate	0.25	0.21	0.27	0.19
Load	2.77	1.36	2.54	1.40
Course tenure	9.63	8.73	10.27	9.60
General tenure	11.41	9.21	12.38	10.3

Notes: Statistics in the table correspond to the course-section level.

Table 2 complements the preceding table by offering score-related data at the course-section level. This perspective is important as it offers information about the variation across sections of a common course as perceived by a student when making enrollment decisions. From the table, the average section size for both Calculus 1 and Calculus 2 is of approximately 33 students. This is consistent with the fact that for Calculus 1, only 27% of the observed sections are operating at their full capacity, whereas this percentage increases to 40% for Calculus 2 (i.e., accreditation requirements place the course capacity constraint at 40 slots). In the second panel of Table 2, the attention shifts to the distribution of scores at the section level. For both Calculus 1 and Calculus 2, approximately 63% of students who enroll in the course successfully complete it. In Calculus 1, the average mean section score is of 2.70 GPA points, while in Calculus 2, it is 2.66. However, a noteworthy aspect to consider is that a substantial proportion of students, accounting for 25% in the case of Calculus 1 and 27% for Calculus 2, opt to drop the section they initially enrolled. This choice to withdraw from a section results in these students not receiving a final score. Subsequent sections deal with the identification challenge resulting from the truncation in the distributions of scores arising from this institutional feature.

In terms of the distribution of instructor characteristics across sections, on average an instructor responsible for a Calculus 1 section is subject to a teaching load of 2.77 courses, while for Calculus 2 the number is 2.54. Regarding course-specific and general tenure, the average Calculus section is taught by a professor with approximately 9.63 terms of experience for Calculus 1 and 10.27 terms for Calculus 2. Additionally, when considering general tenure, the corresponding figures are 11.41 and 12.38 terms.

4 Stylized facts and reduced form exercises

This sections presents some key observations that highlight the variation in the data to be explored in the main empirical analysis. I emphasize three central aspects that will later underpin the construction of the model. First, there is substantial variation in the distribution of scores reported by instructors. This is the case both across professors and for any given professor considered across different academic periods. Second, the returns to increasing student ability on different course outcomes varies across the different instructors. In particular, the ability returns to both current and subsequent courses' scores depend on the identity of the instructor. As a final observation, I illustrate how sections differ in terms of the ability distribution of students who demand them, suggesting that students may select courses based on characteristics correlated with their initial ability.

Variation in scoring outputs across instructors/periods

INTEC operates without an explicit policy requiring instructors to meet a target score distribution within each section. A question that might arise is whether an implicit rule regarding the distribution of scores within each section is followed by Calculus 1 and Calculus 2 instructors. The distinction is relevant for modeling purposes, as constructing estimates on the quantities of interest based on the wrong grading model could lead to bias. Let's inspect the data in two different ways regarding this concern.

The first exercise consists of comparing instructors in terms of the distribution of scores they assign to their students. Under professor specific absolute grading policies, these distributions such differ. On the contrary, a situation in which the differences in these distributions are minimal would suggest some form of implicit target distribution exists. Figure 1 shows this for the 10 professors with the highest number of enrolled students in both Calculus 1 and Calculus 2. In both panels each row is associated with a given instructor, while the columns correspond to the letter score obtained by the student[1]. We can understand a row as the letter score distribution of all students who enroll a section of Calculus 1 under the instructor represented by the row. Both panels suggest important differences in the distribution of scores across instructors. Consider, for instance, the first column corresponding to the fraction of students who obtain a score of A under each instructor. This ranges from 71% to 7%. Similar large differences can be identified for other segments of the letter score distribution and for Calculus 2.

[1]The Figure pools letter scores into 4 categories: $\{A\}$, $\{B+, B\}$, $\{C+, C\}$, and $\{D, F, R\}$.

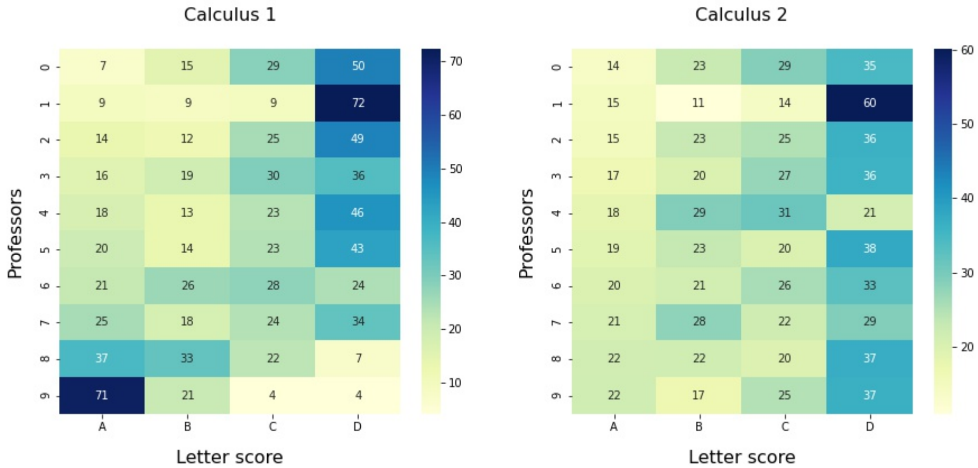


Figure 1: Comparing the distribution of scores across instructors.

Figure 1 is not consistent with a common distribution target across all instructors. In the appendix we show that this is robust to considering differences across instructors withing a common academic term, ruling out concerns of a target distribution that changes across terms. Nevertheless, the evidence above might still be consistent with other grading policies different from absolute grading. In particular it is possible for professors to follow a target distribution for their own sections. Notice that this is different from an absolute grading policy as under the latter the instructor seeks to achieve a given distribution as opposed to setting upfront the way in which the course is to be graded and accepting whatever distribution results from this.

Figure 2 considers the latter possibility by depicting, for different fixed Calculus 1 professors, the distribution of letter scores associated to multiple academic periods. The analog Figure for Calculus 2 is reported in the appendix. Let's consider the four instructors with the highest number of enrolled students, and four each let's show the distribution for the first five academic terms in which they instructor teaches Calculus 1. Intuitively, no variation in the distribution across academic terms would be consistent with a within-professor target distribution. Again, the information in Figure 2 is inconsistent with such behavior as all instructors exhibit non-trivial differences in the distribution of scores from one period to another.

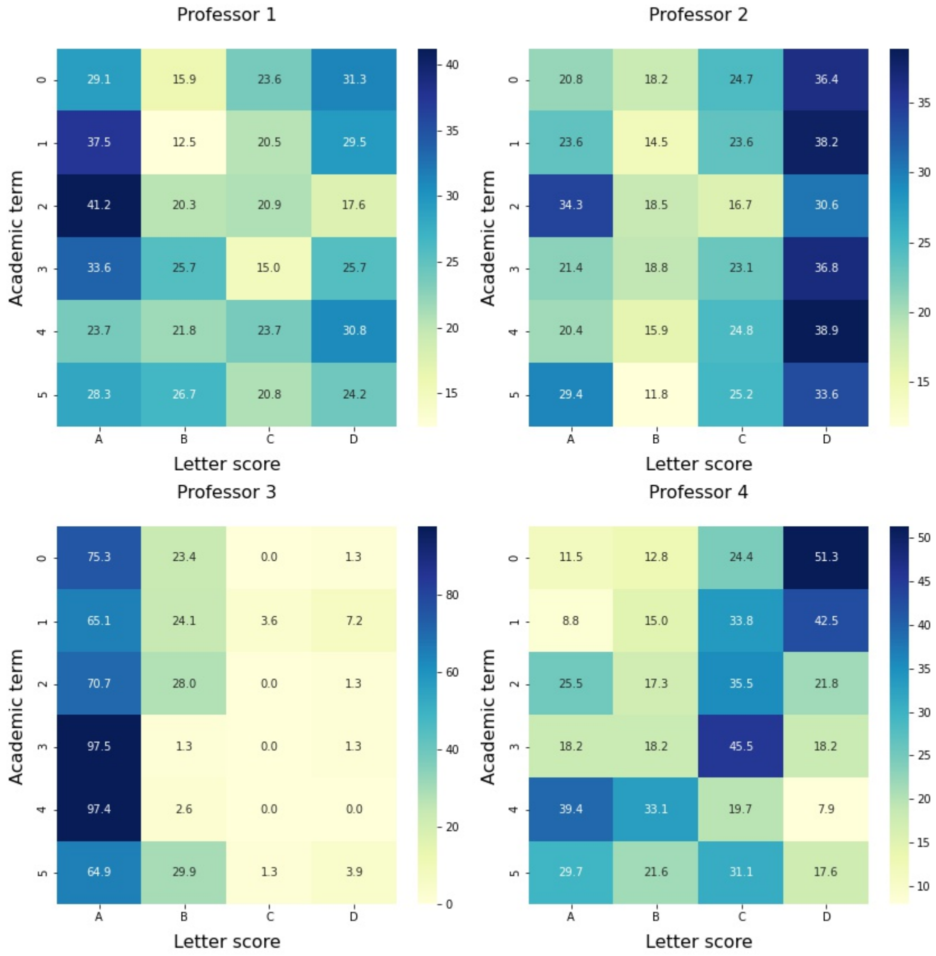


Figure 2: Comparing the distribution of scores across periods.

Score and learning returns to ability differ across instructors.

Central to the main thesis of the paper is the idea that score and learning returns differ across instructors. To start this possibility consider the following simple logistical regression model relating a student's ability to the likelihood of a given discrete learning outcome taking place,

$$y_i = \mathbf{1}\left\{\gamma^0 + \sum_{j\neq 1} \gamma_j^0 \cdot d_{i,j} + \gamma^1 \cdot a_i + \sum_{j\neq 1} \gamma_j^1 \cdot (a_i \times d_{i,j}) + \boldsymbol{\gamma}^2 \mathbf{z}_i + \varepsilon_i \geq 0\right\}.$$

Above, the variable y_i represents a binary outcome, indicating whether a specific learning event has occurred. Let's entertain two specific outcomes for Calculus 1: "student i obtains a score of A" and "student i obtains a fail score". The model above links the likelihood of each of these events to a linear function of the students' characteristics: their ability, denoted as a_i, and a vector of learning-related covariates, represented by $\mathbf{z}_i$. To capture the possibility of professor-specific returns to ability across instructors, consider the dummy variable $d_{i,j}$, for whether student i is matched with professor j, which allow us to write the reduced form model in terms of professor-specific intercepts and ability slopes: γ_j^0, γ_j^1. Of particular interest are differences in the parameters γ_j^1, measuring how an increase in a student's ability affects the likelihood of the event $y_i = 1$ under instructor j (relative to the excluded professor).

Table 3 records the ability slope estimates of the model for the ten instructors with the highest number of enrolled students. The estimates are constructed using only information on first time students as a way of mitigating potential selection concerns. The bottom panel shows the p-value associated with the null hypothesis of all instructors coinciding in terms of their slope parameters (i.e., $\gamma_j^1 = 0$ for all professors other than the excluded instructor).

Table 3: Do score returns differ across instructors?

	A	**Fail/Retire**
Point estimates for the ability slopes (γ_j^1)		
Professor 1 ($\gamma_{j_1}^1$)	-1.14	1.28
	(0.36)	(0.26)
Professor 2 ($\gamma_{j_2}^1$)	-0.47	0.52
	(0.39)	(0.26)
Professor 3 ($\gamma_{j_3}^1$)	-0.40	0.05
	(0.36)	(0.24)
Professor 4 ($\gamma_{j_4}^1$)	-0.99	0.52
	(0.33)	(0.22)
Professor 5 ($\gamma_{j_5}^1$)	-0.83	0.73
	(0.52)	(0.33)
Professor 6 ($\gamma_{j_6}^1$)	0.46	-0.28
	(0.43)	(0.29)
Professor 7 ($\gamma_{j_7}^1$)	-1.24	1.19
	(0.37)	(0.26)
Professor 8 ($\gamma_{j_8}^1$)	-1.92	1.70
	(0.35)	(0.45)
Professor 9 ($\gamma_{j_9}^1$)	-1.91	1.35
	(0.40)	(0.49)
H_0: Joint insignificance of ability slopes (γ_j^1)		
p.value	0.00	0.00

Notes: The ability slope estimates correspond to deviations from the ability slope under the excluded instructor.

For both cases we can reject the null hypothesis of all ability slopes γ_j^1 being simultaneously equal to zero (i.e., the marginal return to ability of all instructors coinciding). The size of the coefficients also suggests nontrivial differences in how increases in the ability level change the likelihood of the learning outcomes. Consider for instance the case of the outcome "i obtains a score A". At the top of the distribution of marginal returns is the sixth professor with a slope of 0.46 (i.e., 0.46 above the excluded professor). This implies that starting from a student for whom the probability of the event is 50% under the excluded professor, a change to instructor j_6 leads to an increase in the probability of the event to approximately 61%. An analog reasoning can be used to argue for large differences for professors at the bottom of the distribution.

One might be concerned about the estimates above reflecting both learning and grading differences. In other words, the differences documented above cannot be interpreted only in terms of learning. The following thought experiment suggests an alternative exercise that deals with this issue: Imagine two students with the same entrance exam score. Enroll both of them with different Calculus 1 instructors, and after completing Calculus 1, have both students enroll under a common Calculus 2 instructor. Under the "ceteris paribus" assumption, since the only distinguishing factor in the paths these students follow is their Calculus 1 instructor, any differences in their Calculus 2 performance should reflect disparities in the learning outcomes associated with their respective Calculus 1 instructors. While the experiment doesn't allow us to directly quantify the magnitude of this gap, as it is influenced by the Calculus 2 professor's own contribution to learning, it can be used to test whether the returns to learning outcomes of the two Calculus 1 instructors differ.

We can frame the experiment above in terms of the reduced form model described below. This mimics the model above except for two things: (i) we now focus on students who enroll a specific Calculus 2 instructor in their second academic period, and (ii) we take y_i as a Calculus 2 learning outcome. As we construct the exercise by fixing the Calculus 2 instructor, we can interpret differences in the ability slopes as resulting from differences in student learning across Calculus 1 instructors.

$$y_i = \mathbf{1}\left\{\gamma^0 + \sum_{j_1=1} \gamma^0_{j_1} \cdot d_{i,j_1} + \gamma^1 \cdot a_i + \sum_{j_1=1} \gamma^1_{j_1} \cdot (a_i \times d_{i,j_1}) + \varepsilon_i \geq 0\right\}.$$

The table presented here resembles the one in the previous exercise. It showcases the point estimates and standard errors for the interaction terms $\gamma^1_{j_1}$. As before, the bottom panel reports the results of a joint test for all ability slopes being equal. We reject the null for the outcome "i obtains a score of A" at a 10% confidence level and at a 5% confidence level for the outcome "i obtains a score of Fail".

Table 4: Do learning returns differ across instructors?

	A	Fail/Retire
Point estimates for the ability slopes (γ_j^1)		
Professor 1 (γ_j^1)	0.71	1.41
	(0.73)	(0.64)
Professor 2 (γ_j^1)	-0.14	1.18
	(0.59)	(0.56)
Professor 3 (γ_j^1)	-0.39	1.99
	(0.76)	(0.85)
Professor 4 (γ_j^1)	-1.04	1.82
	(0.65)	(0.62)
H_0: Equality of the ability slopes (γ_j^1)		
p.value	0.07	0.02

Notes: The ability slope estimates correspond to deviations from the ability slope under the excluded instructor.

Student ability predicts the demand for sections

After their first academic term students can choose which sections of a course to enroll by participating in the course enrollment mechanism. That student characteristics, in particular their ability level, matters for such demand decisions can be seen from the distribution of initial abilities conditional on the professor being enrolled.

Figure 3 depicts these distributions for Calculus 1 in the first panel and Calculus 2 in the second panel. The first panel considers only students who are retaking Calculus 1 and who can therefore participate of the course-enrollment mechanism. Conversely, the second panel corresponds to all students enrolling in Calculus 2 as by construction they all participate in the course-enrollment mechanism. Each row corresponds to one of the ten professors with the highest number of enrolled students, while columns correspond to quartiles in the distribution of student ability with 1 being the lowest quartile. For neither course the distribution of initial abilities is constant across instructors.

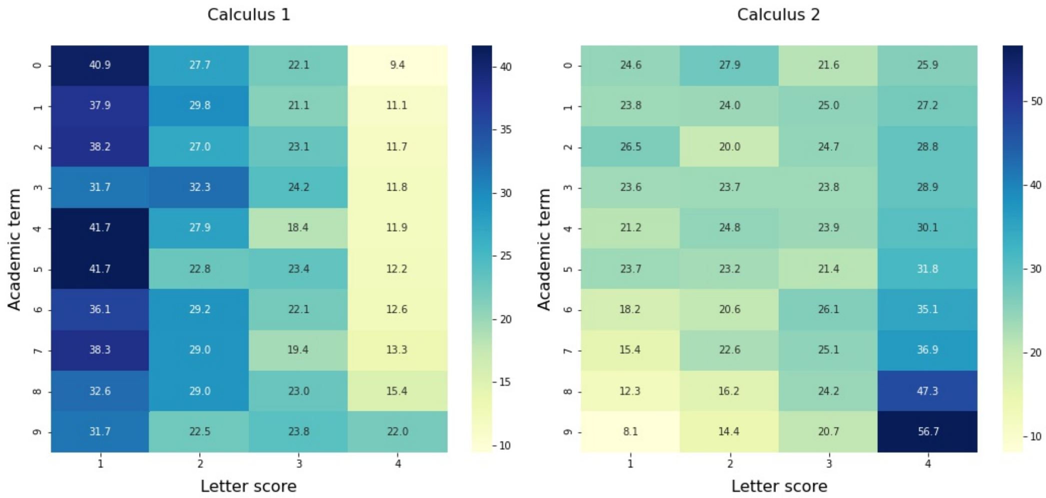


Figure 3: Ability distribution of students who enroll a given section.

5 The model

This section introduces a model that describes a student's academic outcomes along a sequence of compulsory subject-related courses. The model is divided into two key components. First, we outline the process through which learning takes place along the sequence, recognizing both the cumulative nature of learning, and how it depends on the interaction of both student and instructor inputs. Second, we model how students enroll in sections of a course, given the institutional constraints governing our empirical context. The aim is to construct a conceptual version of the model and to leave the formulation of an empirical version to subsequent sections.

5.1 The learning production function

Consider a university that, in each academic term, denoted by $t \in \mathcal{T}$, faces the task of assigning students to instructors leading specific sections of courses in which students seek enrollment. We index an arbitrary student by $i \in \mathcal{I} \equiv \{1, ..., N\}$, and an instructor by $j \in \mathcal{J} \equiv \{1, ..., J\}$. The focus is placed on a sequence of courses centered around a common subject. These courses, which we denote by $\kappa \in \mathcal{K} \equiv \{1, 2, \ldots, K\}$, might correspond for instance to Calculus 1, Calculus 2, Calculus 3, and so forth. Students are required to enroll and successfully complete all of the courses in the sequence in the order specified by the indexes in $\mathcal{K}$. For instance, a section of course $\kappa > 1$ can be enrolled only after obtaining a pass score for course $\kappa - 1$ (e.g., achieving a pass score in Calculus 1 is a prerequisite for enrolling in Calculus 2). This sequential arrangement reflects the curriculum constraints set by the university.

Let $t_i \in \mathcal{T}$ represent the academic period in which student i enrolls in the university. Upon enrollment, i draws an ability, denoted as $a_{i,0} \in \mathbb{R}+$, which we occasionally refer to as i's initial ability type. The value $a_{i,0}$ can be interpreted as i's understanding of the prerequisite material required for the courses in $\mathcal{K}$. For instance, it may correspond to the student's score in the math component of a college entrance examination designed to assess a student's understanding of high-school pre-calculus. As the student progresses along the sequence of courses, its ability is updated in a way that reflects i's acquired knowledge of the sequence curriculum. We denote a student's type at the end of period $t > 0$ as $a_{i,t} \in \mathbb{R}_+$.

In any given academic term t, multiple sections of a course κ may be offered, with each section being guided by a single instructor. The pool of all such instructors is denoted by $\mathcal{J}_t^\kappa \subseteq \mathcal{J}$. Notice that this is potentially a strict subset of $\mathcal{J}$ as some instructors might not be active in certain periods for exogenous reasons. The multiplicity of instructors under a common course implies that more than one way of matching students to professors will exist in any given course/period pair. Let $\kappa_{i,t}$ stand for the course in the sequence student i seeks to enroll in period t, and $j_{i,t}$ for the instructor student i is paired with. While subsequent subsections describe the process by which these assignments take place, our interest here is in describing the academic outcomes conditional on the student's match.

With this goal in mind, let's consider a student, denoted as i, who, in period t, is paired with instructor $j_{i,t} = j$ for course $\kappa_{i,t} = \kappa$. Two potential academic outcomes might arise. First, i might decide to drop the section of the course, in which case an R (i.e., the notation represents 'retire') score is recorded. Such a situation is considered an unsuccessful attempt at completing the course, requiring i to enroll in the same course again in a subsequent term. The dummy variable $R_{i,t}^\kappa$ records i's decision not to drop the section of course κ (i.e., $R_{i,t}^\kappa = 0$ corresponds to dropping the section). Alternatively, the student might choose to complete the course, resulting in a discrete course score (analogous to the A, B+, B, and so forth system common in higher education institutions). Student i's discrete score upon completing the course is recorded by the discrete random variable $S_{i,t}^\kappa$. The setting described above is formally captured by the following collection of equations,

$$
\begin{aligned}
&[\,0\,] \quad && a_{i,t_i} && \sim F_a(\cdot\,), \quad j_{i,t} = j, \quad \kappa_{i,t} = \kappa, \\
&[\,1\,] \quad && a_{i,t} && = f_j(a_{i,t-1}, \mathbf{x}_{i,j,t}), \\
&[\,2\,] \quad && s_{i,t} && = \beta_j \cdot a_{i,t} + c_j, \\
&[\,3\,] \quad && R_{i,t}^\kappa && = \mathbf{1}\big\{s_{i,t} + \tilde{\varepsilon}_{i,j,t}^\kappa \geq s_{l^*}\big\}, \\
&[\,4\,] \quad && S_{i,t}^\kappa && = \sum_l s_l \cdot \mathbf{1}\big\{s_{l+1} > s_{i,t} + \tilde{\eta}_{i,j,t}^\kappa \geq s_l\big\}.
\end{aligned}
$$

To fix ideas, suppose a student i enters academic term t with an ability type given by $a_{i,t-1}$. Equation [1] describes the learning output of such a student after being paired with instructor j for course κ. This quantity, unobserved by the researcher, is denoted by $f_j(a_{i,t-1}, \mathbf{x}_{i,j,t})$. Notice that besides the student's ability, learning outputs are affected by a vector $\mathbf{x}_{i,j,t}$ capturing learning-related covariates. That these learning production functions are indexed by j implies the possibility of different learning outputs across instructors even conditional on the values of $a_{i,t-1}$ and $\mathbf{x}_{i,j,t}$[2]. In turn, equation [2] describes the score outcome the student obtains, $s_{i,t}$. The latter differs from learning in that it is expressed in terms of the grading policy of i's professor, $(\beta j, c_j)$.

Equations [4] describe how i's learning output maps into a course discrete score. Intuitively, we can think of $s_{i,t}$ as the student's expected continuous score obtained in period t. To make it clear that such a score depends on the student's ability and the underlying covariates, we will occasionally use the notation $s_{i,t} \equiv s_j(a_{i,t-1}, \mathbf{x}_{i,j,t})$. Notice that this quantity differs from $S^{\kappa}_{i,t}$, the discrete score obtained by the student. While the former represents the instructor's granular assessment of the student's performance (e.g., the 100 points based raw score i obtains in j's course) the latter is a discrete variable indicating the region of the score support where $s_j(a_{i,t}, \mathbf{x}_{i,j,t})$ falls. Institutional rules fix thresholds $s_1 > s_2 > ... > s_L$ which determine the map between a student's continuous underlying score and its final discrete score. As an example, student i obtains a score of s_l if $s_j(a_{i,t}, \mathbf{x}_{i,j,t})$ (plus a random perturbation) exceeds the threshold s_l but falls short of the threshold for score s_{l+1}. The error terms $\tilde{\eta}^{\kappa}_{i,j,t}$ and $\tilde{\varepsilon}^{\kappa}_{i,j,t}$ perturb the relationship between a student's continuous and discrete course scores.

We highlight that $s_{i,t}$ depends not only on the learning generated by the match but also on the instructor's grading policy (β_j, c_j). We can interpret these as encapsulating the leniency or stringency with which a student's learning is evaluated in the course. Figure 4 depicts this by plotting the map $a_{i,t} \rightarrow \beta_j \cdot a_{i,t} + c_j$ for two different instructors who differ only in their grading policies (i.e., but whose learning production functions coincide: $f_j = f_{j'}$). For instance, the blue curve depicts an instructor who, while more lenient in terms of the marginal return to learning (e.g., a higher β_j), is more stringent in terms of the level of the scoring equation (e.g., a smaller c_j). These differences map two students, with the same underlying learning output, to different scores under each of the professors. For example, while under the red curve students with low ability levels end up above the s_l threshold, the same is not true under the scoring equation corresponding to the blue curve.

[2]This can be interpreted in terms of allowing for pedagogical differences across instructors which are not captured by any of the measured inputs of the learning production functions.

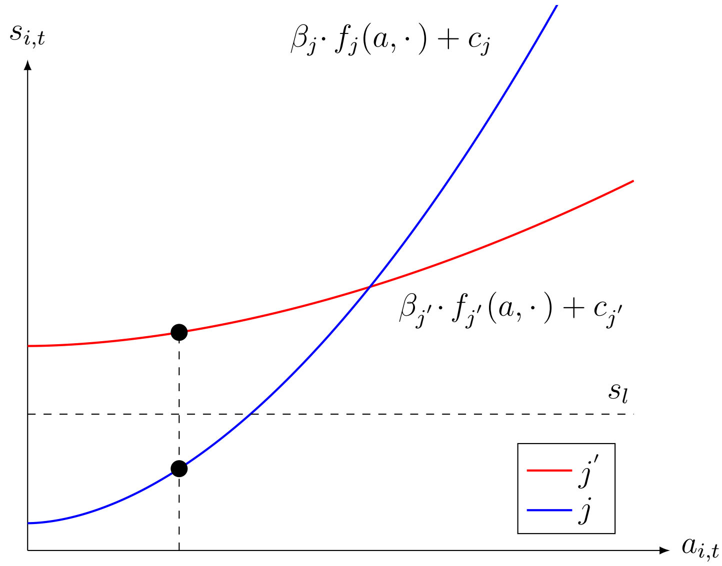


Figure 4: Grading policy differences

As explained before, student i can choose to drop instructor j's section of course κ which is explained by equation [3]. Intuitively, i chooses to drop the section whenever its underlying continuous score, $s_j(a_{i,t}, \mathbf{x}_{i,j,t})$, places him below a certain threshold s_{l^*}. As an example, one can think of a student choosing to drop the course whenever it expects to end up with a fail score. Notice that the error term $\tilde{\varepsilon}^{\kappa}_{i,j,t}$ allows for heterogeneity in the course dropping threshold. It can also capture uncertainty resulting from course dropping decisions depending on noisy signals of the true underlying score (i.e., the student's perception of its position in the grading policy after the midterm, but without the final exam signal). An error term $\tilde{\varepsilon}^{\kappa}_{i,j,t}$ with a large variance could for instance correspond to a situation in which students face a lot of uncertainty before the dropout deadline. Correlations between $\tilde{\varepsilon}^{\kappa}_{i,j,t}$ and $\tilde{\eta}^{\kappa}_{i,j,t}$ describe unobserved relationships between the scoring and course dropping outcomes[3].

We conclude our description by being precise about the units under which learning is being measured. In a setting described by standardized testing, all students are tested under a common grading policy so that a natural choice is to measure $f_j(a, \mathbf{x})$ in the units of the common test. This is not the situation in our empirical context as in higher education institutions students are evaluated according to the grading policy of their matched instructor. We instead propose defining learning output in terms of the grading policy of a reference professor $\hat{j}^{\kappa}$ for each course κ[4]. It then follows that for any course κ instructor j, we interpret $f_j(a, \mathbf{x})$ as the learning output of a student who is instructed by professor j but graded according to instructor $\hat{j}^{\kappa}$'s grading policy. We can

[3]Following the midterm/final analogy, we can interpret correlations between the error terms as capturing the fact that the student's position after the midterm is a noisy signal of what its final continuous score will be. In other words, a student who outperforms its own type after the midterm is also likely to do so for the final exam.

[4]The identity of the reference professor is irrelevant and any course κ professor can serve this role.

then interpret (β_j and c_j) as deviations of instructor j's grading policy from that of the reference professor. Under this intuition, c_j corresponds to the baseline score granted by the instructor and β_j captures the marginal reward to learning under instructor j (i.e., in both cases relative to the reference professor). It is immediate that under the proposed normalization, $\beta_{\hat{j}^\kappa} = 1$ and $c_{\hat{j}^\kappa} = 0$.

5.2 The demand for sections within a course

The preceding subsection provides a model for how learning and related academic outcomes are determined given a student-instructor match. Now, we describe how these matches emerge in our empirical setting. Two rules govern the assignment of students to professors at Intec. First, all first-period students (i.e., students in their initial academic period) are randomly assigned to a section of course $\kappa = 1$. Second, all other course enrollment instances, require the student to enroll a section of a course κ by participating in a first-come-first-serve mechanism. To be precise, every academic term t a course enrollment platform will open enabling students to enroll in a section of the course. Since multiple sections can be associated with the same instructor, we must introduce additional notation that distinguishes two sections under the same instructor. In particular, consider denoting a particular section as $s \in Sec_t^\kappa$, where Sec_t^κ represents the collection of all sections of course κ active in period t. Of course, each of these sections must be under an instructor j in $\mathcal{J}^\kappa t$. Whenever it is not obvious from the context, we explicitly keep track of the professor associated with section s using the notation j_s.

The course-enrollment mechanism implies that a student attempting to enroll in a section of course κ faces two sequential decisions. First, student i must choose an entry time, $\tau > 0$, to access the platform. Subsequently, i must select, from the available sections, which one to enroll in. These two problems are intertwined since sections are subject to capacity constraints (i.e., limited slots are available for each section due to institutional constraints). This implies that students may need to access the platform early to secure enrollment in highly demanded sections. Formally, we frame the decision problem of a student i seeking a section of course κ in terms of the following two-stage optimization problem,

$$\max_{\tau \geq 0} \left[\max_{s} \big\{ U_{i,s,t} \text{ s.t. } s \in \mathcal{C}_t(\tau) \big\} + \phi(\tau) \right],$$
$$\mathcal{C}_t(\tau) \equiv \{ s \in Sect_t^\kappa : \tau \leq \tau_{s,t}^{eq} \}.$$

The inner maximization problem corresponds to a standard discrete choice problem, where students choose the section of the course that maximizes their utility, denoted by $U_{i,s,t}$. Importantly, students can only select sections from the set $\mathcal{C}_t(\tau)$, which includes all active sections whose capacity constraint is not binding at the entry time τ. In other words, i's choice set in period t may be potentially smaller than $Sect_t^{\kappa}$, the set of all active sections in t. Since the availability of a slot in a given section depends on the demand decisions of other students, we must treat the choice set faced by i as an equilibrium object. The term $\tau_{s,t}^{eq} > 0$, assumed to be known by the students, denotes the equilibrium time at which the capacity constraint of section s becomes binding. For instance, very popular sections will be associated with small values of $\tau_{s,t}^{eq}$, while the opposite holds for unpopular sections.

In turn, the outer maximization problem pertains to the decision of when to enter the platform, while recognizing that this choice influences the set of options the student will ultimately face. The formulation above assumes that students face a cost from participating in the course enrollment mechanism, $\phi(\tau)$, and that such a cost is a function of their platform entry time decision. This cost rationalizes the fact that not all students choose to enter the platform at the same time and can be interpreted as a reluctance towards early enrollment or more generally of participating in the mechanism.

We adopt a random utility model approach for the inner maximization problem by treating $U_{i,s,t}$ as a random variable. This allows us to model preferences in terms of a systematic component, shared by all students with common characteristics, as well as an idiosyncratic component capturing unobserved heterogeneity in students' preferences. In particular, as seems reasonable from our descriptive evidence exercises, we assume student i's utility for section s under an instructor j takes the following form,

$$U_{i,s,t} = U_{s,t}(s_j(a_{i,t}, \mathbf{x}_{i,j,t})\ ,\ f_j(a_{i,t-1}, \mathbf{x}_{i,j,t})) + \nu_{i,s,t}.$$

Intuitively, our model for section preferences postulates that students derive utility not only from the score they expect to obtain under instructor j but also from the actual learning derived from the match. Different functional forms for $U_{s,t}(\cdot)$ can be used to capture various preferences for these two components within the student population. For example, at the extremes, students might have preferences that depend on only scores or learning. As suggested by the indexing of the systematic utility, students might also have preferences related to other aspects of the section being demanded, such as the course schedule or characteristics of the instructor, not directly related to the learning or scoring outcomes expected by the student. These preferences can be incorporated into the formulation above by using, for example, preference fixed effects as part of the systematic utility $U_{s,t}$ specification. The term $\nu_{i,s,t}$ is an error term reflecting the idiosyncratic component of utility.

While the above formulation clearly outlines the two steps involved in the course demand problem faced by students, it is also possible (and potentially advantageous from an empirical point of view) to express the demand problem from a different perspective.

Namely, one can think of students first choosing a section s from the full set of active sections $Sect_t^{\kappa}$, and subsequently choosing a platform entry time that maximizes their utility conditional on securing a slot at the section choice. We can derive this alternative formulation by manipulating the expression for the demand model as in the following,

$$\max_{\tau}\left[\max_{s}\left\{U_{i,s,t} \text{ s.t. } s \in \mathcal{C}_t(\tau)\right\} + \phi(\tau)\right] = \max_{s}\left[U_{i,s,t} + \max_{\tau}\left\{\phi(\tau); \text{ s.t. } s \in \mathcal{C}_t(\tau)\right\}\right].$$

One can think of this formulation as the decision of student i from an ex-ante perspective. Before the platform opens the student face no constraints in its choice set, as it can always choose to enter the platform sufficiently early (i.e., which requires accepting the cost of such decision) in a way that ensures the availability of a slot in the section being demanded.

Discussion

In this project, we consider the possibility of suboptimal student-instructor assignments emerging due to institutional policies governing how these matches take place. Having introduced the key aspects of our framework, we can now inquire about how such undesirable allocations might arise in our setting. The first channel is quite obvious: in the presence of matching effects, the random assignment of first-term students to sections with $\kappa = 1$ is unlikely to result in an optimally learning assignment. In the following discussion, we emphasize a second channel: students, apart from considerations regarding learning, have preferences for the scores they expect to achieve under a potential instructor match. If instructors associated with high learning outputs differ from those associated with lenient grading policies, some students at the margin might end up demanding suboptimal instructors, from a learning perspective. A toy example illustrates the mechanism behind this argument.

Consider for instance a scenario where the learning production functions take the following form, $f_j(a_i) = \delta_j \cdot a_i$. We entertain the problem of assigning two students, i_H and i_L, to two instructors, j_H and j_L (each under a single slot section). We assume that these students have initial abilities described by $a_{i_H} > a_{i_L}$, and that $\delta_{j_H} > \delta_{j_L}$ so that both students are more learning-productive under j_H relative to j_L. For the sake of simplicity, let's also set aside the heterogeneity and uncertainty in the learning and scoring equations by assuming the $\tilde{\eta}$ and $\tilde{\varepsilon}$ terms to be exactly equal to zero. Two assignments are possible, either (j_H, i_H) or (j_H, i_L), with the leading to a higher aggregate learning output than the latter. We are concerned about the question of whether or not such assignment would arise under the observed course enrollment mechanism.

It is useful to think about the question under the extreme scenarii of students who care only about scores or only about learning. For instance, in the latter case a course enrollment mechanism as the one in our setting must lead to the ideal assignment. To see this, suppose that this is not the case so that we observe the match (i_L, j_H) in the data. Since in our model students freely demand course sections at the equilibrium costs $\phi_j \equiv \phi(\tau^{eq}_{\kappa,j})$, our observation implies that $f_{j_H}(a_{i_L}) + \phi_{j_H} > f_{j_L}(a_{i_L}) + \phi_{j_L}$ or equivalently $f_{j_H}(a_{i_L}) - f_{j_L}(a_{i_L}) > \phi_{j_L} - \phi_{j_H}$. In other words, i_L is willing to assume the equilibrium cost associated to demanding j_H's section. Notice however that given the nature of the learning production function, it must also be the case that i_H finds it optimal to demand such section since (i.e., $f_{j_H}(a_{i_H}) - f_{j_L}(a_{i_H}) > f_{j_H}(a_{i_L}) - f_{j_L}(a_{i_L})$). Since at equilibrium this cannot be the case, the equilibrium cost will adjust until student i_L is discouraged from demanding j_H. The opposite situation could arise when students are only concerned about their score outcomes: $\beta_j f_j(\cdot) + c_j$. In this case it is easy to construct grading policy conditions under which the demand patterns described above are reversed. For instance, if $\beta_{j_H} \cdot \delta_{j_H} < \beta_{j_L} \cdot \delta_{j_L}$, while both students would prefer instructor j_L, student i_H ends up matched to such an instructor as it can outbid i_L under the current assignment mechanism.

This example illustrates one of the main tensions we will be exploring in subsequent sections. It also highlights the fact that the outcome assignment depends on the course enrollment mechanism used by the university. For instance, in our example a priority mechanisms that allows certain students to enroll sections first might lead to an either desirable or undesirable assignment depending on the context considered, by reducing the competition faced by certain students when demanding slots in a course. The complexity associated to studying any of these situations theoretically suggests an empirical approach is better suited for such purposes. In subsequent sections we follow this proposal by estimating the primitives of the model and conducting counterfactual exercises under different assignment mechanisms a university's administration might entertain.

6 Empirical model and identification arguments

In this section, we explore the identification of an empirical version of the model described above. Before delving into this, it is beneficial to compare our framework with other commonly utilized empirical models for quantifying disparities in pedagogy among instructors. This comparison will facilitate the placement of our model within the existing literature and underscore certain identification challenges arising from our divergences from these established models. To illustrate, let's consider the following model for generating learning outcomes presented below[5].

[5] We use the same notation as in the model described in the preceding section.

$$
\begin{aligned}
&[\,0\,] \qquad a_{i,0} \sim F_a(\cdot\,), \\
&[\,1\,] \qquad s_{i,1} = f_j(a_{i,0}) + \tilde{\eta}_{i,j,1}.
\end{aligned}
$$

This simple model captures summarizes (in essence) a substantial body of work in assessing pedagogical disparities among instructors. For instance, assuming $f_j(\cdot\,)$ is additive separable in student and professor attributes is a common feature in many educational studies. As another example, the fact that scores are expressed in the same units as the learning production function corresponds to situations in which standardized tests are available. The prevalence of such a model in the literature largely stems from its high tractability from an econometric perspective, enabling the estimation of learning returns associated with each instructor by directly analyzing within-professor score distributions. For instance, in the case above, the average score for students enrolled under instructor j given the ability type a_0 serves as a consistent estimator for the functions $f_j(a_0)$.

As is clear from our previous discussions, the model is a poor description of higher education environments, which is why we choose to deviate from it. Nevertheless, each of these deviations presents empirical challenges that render the identification approach described earlier inapplicable. Let's see this by means of some examples. To be concrete, consider a minor modification of the previous model as to account for differences in instructors' grading policies while keeping other aspects the same,

$$
\begin{aligned}
&[\,0\,] \qquad a_{i,0} \sim F_a(\cdot\,), \\
&[\,1\,] \qquad s_{i,1} = \beta_j \cdot f_j(a_{i,0}) + c_j + \tilde{\eta}_{i,j,1}.
\end{aligned}
$$

Even without considering the other elements in our framework, it is evident that the within-professor conditional average approach discussed earlier is no longer useful in identifying the learning production functions. For example, the average scores of students under instructor j conditional on the ability type $a_{i,0} = a_0$, is now consistent for a quantity that conflates both learning returns and grading policies, $\frac{1}{n}\sum_i s_{i,1} \quad \rightarrow_p \quad \beta_j \cdot f_j(a_0) + c_j$. Put simply, observing high average scores may indicate either a high learning return under professor j, a choice of a very lenient grading policy, or both. Clearly, the latter is unsatisfactory if the aim is to deduce the nature of an instructor's production function.

As a second example, consider the following alternative deviation from the model in the direction of our framework. Specifically, let's modify the model by allowing students to withdraw from previously enrolled courses/sections. Following our formulation, an example of this corresponds to the following,

$$
\begin{aligned}
&[\,0\,] \qquad a_{i,0} \sim F_a(\cdot\,),\\
&[\,1\,] \qquad s_{i,1} = f_j(a_{i,0}) + \tilde{\eta}_{i,j,1},\\
&[\,2\,] \qquad R_{i,1} = \mathbf{1}\{s_{i,1} + \tilde{\varepsilon}_{i,j,1} \geq s_{l^*}\}.
\end{aligned}
$$

Since only the scores of students who choose not to withdraw from a course can be observed in the academic records, the approach based on the within-professor average scores of students conditional on ability must also condition on the students not choosing to withdraw from the section of the course. In this case, such an average is again consistent for a quantity that differs from the learning production function images of interest, $\frac{1}{n}\sum_i s_{i,1} \;\; \rightarrow_p \;\; f_j(a_0) + \mathbb{E}(\tilde{\varepsilon}_{i,j,1}\,|\,R_{i,1} = 1)$. The implication is that, after observing high average scores for the conditioning set, the researcher is unable to determine whether these scores reflect a high learning return or merely the fact that the average is computed for students with high $\varepsilon_{i,j,1}$ draws.

The concerns listed above, while not exhaustive, clarify the point. The modeling realism gains resulting from our framework come at the cost of needing additional work to estimate the primitives of interest. In what follows, we consider some arguments in this direction. Mimicking the previous section, we divide the discussion into the identification of the learning production function primitives and the identification of the course/section model model.

6.1 Identifying the learning production functions

Consider the problem of inferring the shape of the learning production function associated to a given instructor in course $\kappa \in \mathcal{K}$. As explained before, the main challenge this poses lies in disentangling the contributions of grading policies and actual learning outputs on the observed distribution of scores. To address this, we exploit the sequential enrollment of students into courses in the sequence $\mathcal{K}$, and the fact that while learning in course κ impacts outcomes in course $\kappa + 1$, the same is not true about the grading policies used by κ instructors. The starting point is a set of assumptions regarding the distribution of the error terms in our model.

Assumption 1. *The following assumptions are assumed to hold,*

1. *Random variables $\eta^{\kappa}_{i,j,t}$ and $\varepsilon^{\kappa}_{i,j,t}$ exist such that $\tilde{\varepsilon}^{\kappa}_{i,j,t} = \sigma^{\kappa}_{\varepsilon}\cdot\, \varepsilon^{\kappa}_{i,j,t}$ and $\tilde{\eta}^{\kappa}_{i,j,t} = \sigma^{\kappa}_{\eta}\cdot\, \eta^{\kappa}_{i,j,t} + \sigma^{\kappa}_{\varepsilon}\cdot\, \varepsilon^{\kappa}_{i,j,t}$ for the scalars $\sigma^{\kappa}_{\varepsilon}$,$\sigma^{\kappa}_{\eta}$,*

2. *The sequences $\{\eta^{\kappa}_{i,j,t}\}_{i,j,t}$ and $\{\varepsilon^{\kappa}_{i,j,t}\}_{i,j,t}$ are mean zero and i.i.d.. Their distributions, denoted by $F_{\eta}(\cdot\,)$ and $F_{\varepsilon}(\cdot\,)$, are known by the researcher. The associated densities are denoted by $f_{\eta}(\cdot\,)$ and $f_{\varepsilon}(\cdot\,)$.*

3. The random variable $\nu_{i,s,t}$ is independent of from $(\eta^{\kappa}_{i,j,t},\ \varepsilon^{\kappa}_{i,j,t})$.

The first two parts of the assumption are technical and are primarily used as tools to facilitate inversion arguments in identifying the learning production functions. Essentially, we assume that the distribution of the error terms in the scoring equation and the course dropping equation can be parameterized in terms of their variances. Correlations between these error terms are integrated into the model through sums of random variables (i.e., $\sigma^{\kappa}_{\eta}\cdot \eta^{\kappa}_{i,j,t} + \sigma^{\kappa}_{\varepsilon}\cdot \varepsilon^{\kappa}_{i,j,t}$ is correlated with $\sigma^{\kappa}_{\eta}\cdot \eta^{\kappa}_{i,j,t}$). The third part of the assumption requires that unobserved heterogeneity in the preferences of students over courses/sections remains unrelated to the perturbations of the scoring and dropping equations. Although the latter involves restrictions, we aim to mitigate potential correlations by incorporating a comprehensive set of controls into our demand specification when conducting our empirical exercises.

Let's now construct an argument for the identification of the learning production function given Assumption 1. For expositional reasons, we present results for a simplified version of our model and leave a treatment of the fully fledged framework for the appendix section. In particular, we consider a version of our model in which students don't have the option of dropping a section of a course. This allows us to bypass some technical details which are not a the core of the results. Second, the focus here is on the identification of the production function primitives associated to instructors in the first course of the sequence, $\kappa = 1$. Constructing arguments for other courses will be a simple matter of adapting the notation in what follows. In addition, since our arguments will not depend on the specific time period in which a student enrolls a course/section but instead just require keeping track of whether a student is in its first or second academic term in the university, we simplify the notation by omitting the time indices. It will be clear from the context whether an argument is based on first or second term students.

With this in mind, consider the problem of identifying the learning production function of a $\kappa = 1$ instructor, j^1. Our analysis focuses on the collection of all students who in their first enrollment instance of course $\kappa = 1$ obtain a score of s_l or higher conditional on enrolling a section under j^1. Furthermore, we condition on students of an initial type $a_{i,t_i} = a_0$ and who enroll j^1's section under a vector of covariates $\mathbf{x}_1$. Our structural model offers an expression for the conditional probability described above.

$$\begin{aligned}
\mathbb{P}\big(S^1_{i,j^1} \geq s_l \ \mid \ a_0,\ \ \mathbf{x}_1,\ \ j^1\big) &= \int_{\eta} \mathbf{1}\big\{\beta_{j^1}\cdot f_{j^1}(a_0,\mathbf{x}_1) + c_{j^1} + \sigma^1_{\eta}\cdot \eta \geq s_l\big\}\ f_{\eta}(\eta)\, d\eta, \\
&= \int_{\eta} \mathbf{1}\bigg\{\eta \geq \frac{s_l - \beta_{j^1}\cdot f_{j^1}(a_0,\mathbf{x}_1) - c_{j^1}}{\sigma^1_{\eta}}\bigg\}\ f_{\eta}(\eta)\, d\eta, \\
&= \bigg[1 - F_{\eta}\bigg(\frac{s_l - \beta_{j^1}\cdot f_{j^1}(a_0,\mathbf{x}_1) - c_{j^1}}{\sigma^1_{\eta}}\bigg)\bigg].
\end{aligned}$$

In words, student i achieves a score above s_l whenever $\beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1} + \sigma^1_\eta \cdot \eta^1_{i,j^1}$ (i.e., the students expected continuous score) falls weakly above the threshold s_l. The expression above just establishes a relationship between the observed mass of students satisfying the event $S^1_{i,j^1} \geq s_l$ (within the conditioning set), and a function of the primitives of the model. Under Assumption 1, we can invert the relationship to obtain the equivalent expression given below,

$$\frac{s_l - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\eta} = \underbrace{F^{-1}_\eta \Big[\mathbb{P}\big(S^1_{i,j^1} \geq s_l \mid a_0, \ \mathbf{x}_1, \ j^1\big) \Big]}_{\equiv \theta(s_l \mid a_0, \mathbf{x}_1, j^1)}.$$

The latter lends itself to an intuitive interpretation. Consider the marginal student, whose η^1_{i,j^1} draw places him precisely at the boundary between scores s_l and s_{l-1}. Within the conditioning set, such marginal student's η draw defines the entire mass of students who ultimately receive a score above s_l within the conditioning set. For example, those with a higher η^1_{i,j^1} will achieve scores weakly above s_l, while those with smaller draws obtain a strictly smaller score. The right-hand side of the expression above identifies the marginal student's draw by finding the precise η^1 value such that the mass to its right under $F_\eta(\cdot)$ corresponds exactly to the observed share of students who obtain a score above s_l, $\mathbb{P}\big(S^1_{i,j^1} \geq s_l \mid a_0, \ \mathbf{x}_1, \ j^1\big)$ (i.e., which the econometrician can observe). Importantly, we can also identify the marginal student as that with a draw satisfying $(s_l - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1} + \sigma^1_\eta)/\sigma^1_\eta = \eta$. The equality derived merely states that these two expressions identifying the marginal student must coincide. The notation $\theta(s_l \mid a_0, \mathbf{x}_1, j^1)$ makes it clear that any change in the conditioning arguments leads to a change in the marginal student's identity.

By itself $\theta(s_l \mid a_0, \mathbf{x}_1, j^1)$ does not offer much insight into the underlying model as it pools multiple primitives into a single expression. However, as stated in the following proposition, when considered for two different score cutoffs, s_l and $s_{l'}$, it is possible to start gaining an understanding of some primitives of interest.

Proposition 1. *The images $f_{\hat{j}^1}(a_0, \boldsymbol{x}_1)$ (i.e., $\kappa = 1$'s reference learning production function) and the variance parameter σ^1_η are point identified.*

Proof. Fixing the conditioning quantities $a_0, \mathbf{x}_1$, we can identify the marginal students associate to the letter scores s_l and $s_{l'}$,

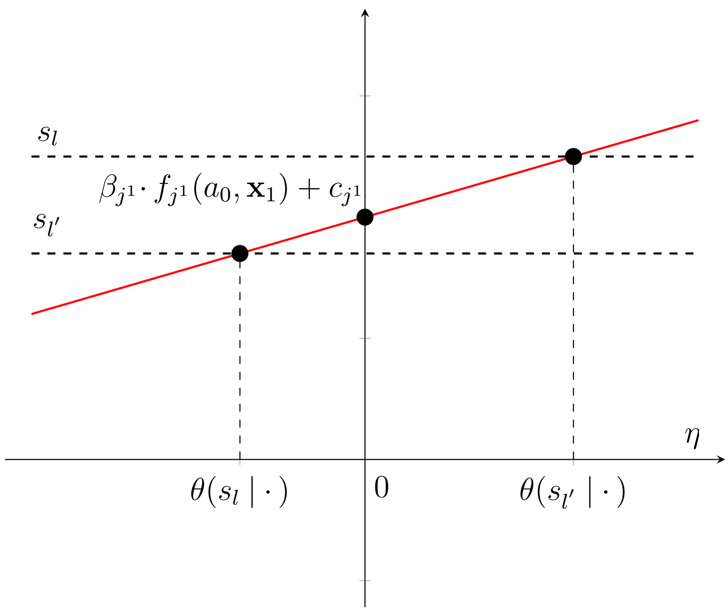
$$\theta(s_l|a_0, \mathbf{x}_1, j^1) = \frac{s_l - \beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\eta} \text{ and } \theta(s_{l'}|a_0, \mathbf{x}_1, j^1) = \frac{s_{l'} - \beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\eta}.$$

When $l \neq l'$, the latter defines a system of two equations on the unknowns σ^1_η and $\beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1}$. Solving for the unique solution to the system leads to the following expression,

$$\sigma_\eta^1 = \frac{s_l - s_{l'}}{\theta(s_l \mid a_0, \mathbf{x}_1, j^1) - \theta(s_{l'} \mid a_0, \mathbf{x}_1, j^1)},$$
$$\beta_{j^1} f_{j^1}(a_0) + c_{j^1} = s_l - \frac{s_l - s_{l'}}{\theta(s_l \mid a_0, \mathbf{x}_1, j^1) - \theta(s_{l'} \mid a_0, \mathbf{x}_1, j^1)} \cdot \theta(s_l \mid a_0, \mathbf{x}_1, j^1).$$

The identification of the image $f_{\hat{j}^1}(a_0, \mathbf{x}_1)$ follows from considering the second equation above for $j^1 = \hat{j}^1$ and recalling that for the reference instructor $\beta_{\hat{j}^1} = 1$ and $c_{\hat{j}^1} = 0$. ■

A graph serves as a visual representation of the argument behind the proof. Within the conditioning set, we can think of a student i's score as a linear function of its unobserved draw η_{i,j^1}^1, with an intercept of $\beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1}$ and a slope of σ_η^1. Our identification of the marginal student associated to each letter score corresponds to identifying a point in the score equation. In the graph, for instance, our arguments allow us to identify the points $(\theta(s_l \mid \cdot), s_l)$ and $(\theta(s_{l'} \mid \cdot), s_{l'})$. The fact that a linear equation is pinned down by two of its points allows us to identify both the slope and the intercept of the curve. In more intuitive terms, the result states that we can always identify the learning output of an instructor j^1, in terms of its own grading policy, by directly inspecting the distribution of scores such instructor induces.



It must be emphasized that in the absence of variations in grading policies among professors, the aforementioned arguments would allow us to fully identify the learning production functions associated to each instructor. The situation resembles a scenario under standardized tests where observed scores directly reflect disparities in teaching abilities. In our context, the presence of grading policies necessitates additional efforts to separate the effects of learning returns from grading policies. With this purpose in mind let's consider the performance of students in our conditioning set in the subsequent course,

$\kappa = 2$. To be precise, we are interested in the fraction of students (within our conditioning set) who after successfully completing $\kappa = 1$, enroll a section of $\kappa = 2$ under instructor j^2 and obtain a score above s_l. In addition, we focus our attention of the subpopulation of students who enroll j^2's section under a vector of covariates $\mathbf{x}_2$. As before, our model implies a concrete expression for the conditional probability of the event described above,

$$
\begin{aligned}
&\mathbb{P}\big(S^2_{i,j^2} \geq s_l, \ \mid a_0, \ \mathbf{x}_1, \ \mathbf{x}_2, \ j^1, \ j^2, \ S^1_{i,j^1} \geq s_{l^*}\big), \\
&\quad = \left[1 - F_\eta\left(\frac{s_{l^*} - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\eta}\right)\right]\left[1 - F_\eta\left(\frac{s_l - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2}}{\sigma^2_\eta}\right)\right], \\
&\quad = \left[1 - F_\eta\Big(\theta(s_{l^*} \mid a_0, \mathbf{x}_1, j^1)\Big)\right]\left[1 - F_\eta\left(\frac{s_l - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2}}{\sigma^2_\eta}\right)\right].
\end{aligned}
$$

This expression bears a close resemblance to the one discussed earlier for the identification of $\theta(s_l \mid a_0, \mathbf{x}_1, j^1)$. The difference lies in the consideration of students who not only achieve a score of s_l or above in course $\kappa = 2$, but also those who are assigned to a specific instructor j^1 in course $\kappa = 1$ and successfully complete the course under such professor. Inverting the relationship we obtain the following identity,

$$
\frac{s_l - \beta_{j_2} \cdot f_{j_2}(f_{j_1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j_2}}{\sigma^2_\eta} = \underbrace{F^{-1}_\eta\left[\frac{\mathbb{P}\big(S^2_{i,j^2} \geq s_l \mid \ a_0, \ \mathbf{x}_1, \ \mathbf{x}_2, \ j^1, \ j^2, \ S^1_{i,j^1} \geq s_{l^*}\big)}{(1 - F_\eta(\theta(s_{l^*} \mid a_0, \mathbf{x}_1, j^1))}\right)}_{\theta(s_l \mid a_0, \mathbf{x}_1, \mathbf{x}_2, j^1, j^2)}.
$$

We are now in a position that allows us to state the main result of this section. The result establishes the identification of the $\kappa = 1$ production functions under an injectivity assumption. We state and prove the result before considering an intuitive discussion of the content behind the Proposition.

Proposition 2. *The following identification results hold,*

1. *The image of the composition $\beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \boldsymbol{x}_1), \boldsymbol{x}_2) + c_{j^2}$ and the variance term σ^2_η are point identified,*

3. *Suppose that $f_{j^2}(\cdot, \boldsymbol{x}_2)$ is injective for $\boldsymbol{x}_2$ fixed. Then the image $f_{j^1}(a_0, \boldsymbol{x}_1)$ is point identified provided the existence of $\tilde{a}_0$ such that $f_{j^2}(f_{j^1}(a_0, \boldsymbol{x}_1), \boldsymbol{x}_2) = f_{j^2}(f_{\hat{j}^1}(\tilde{a}_0, \boldsymbol{x}_1), \boldsymbol{x}_2)$.*

Proof. We start by following the same reasoning as in the previous proposition. In particular, fixing the conditioning variables $(a_0, \mathbf{x}_1, \mathbf{x}_2)$, consider the following system of equations on the unknowns σ_η^2 and $\beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) + c_{j^2}$,

$$\frac{s_l - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2}}{\sigma_\eta^2} = \theta(s_l \mid a_0, \mathbf{x}_1, \mathbf{x}_2, j^1, j^2),$$
$$\frac{s_{l'} - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2}}{\sigma_\eta^2} = \theta(s_{l'} \mid a_0, \mathbf{x}_1, \mathbf{x}_2, j^1, j^2).$$

The first claim follows from noticing that when considering $l \neq l'$, the equations above define a system with a unique solution identifying both σ_η^2 and $\beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) + c_{j^2}$.

Consider now the final claim in the proposition. Under the premise, we can find an ability level $\tilde{a}_0$ such that $\beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) + c_{j^2} = \beta_{j^2} \cdot f_{j^2}(f_{\hat{j}^1}(\tilde{a}_0, \mathbf{x}_1), \mathbf{x}_2) + c_{j^2}$. Moreover, given the validity of the first part of the claim in the proposition, whose truth we have already asserted, we can find $\tilde{a}_0$ by directly inspecting the observed data. It follows from the injectivity of $f_{\hat{j}^2}(\cdot)$ that $f_{j^1}(a_0, \mathbf{x}_1) = f_{\hat{j}^1}(\tilde{a}_0, \mathbf{x}_1)$. However, since Proposition 1 has already established the identification of $\hat{j}_1$'s production function, the latter equality implies we can directly infer the image of j^1's production function at the argument $(a_0, \mathbf{x}_1)$.

■

Proposition 1, our main identification argument, can be understood in terms of a simple though experiment. Suppose we observe two students with the same initial ability level a_0 but assigned to different $\kappa = 1$ instructors: student one with instructor j^1 and student two with $\tilde{j}^1$. As discussed, comparing their $\kappa = 1$ scores directly is uninformative for discerning potential gaps in their learning outputs. The disparity in the scores could be due to either instructor quality differences or differences in the grading policies used by these instructors.

A potential solution to this issue consists in comparing these two students, not in terms of their $\kappa = 1$ scores, but in terms of some other future signal related to $\kappa = 1$'s learning returns but not $\kappa = 1$'s grading policies. Carefully choosing such signals then becomes very important. For instance, one concern is that as we increase the time distance between the enrollment of $\kappa = 1$ and the measurement of the signal, the noise in the latter might increase, making it difficult to detect differences in instruction quality empirically. Moreover, one might be concerned about differences in the academic path followed by the students after $\kappa = 1$ enrollment, and prior to the measurement of the signal, which would invalidate the ideal type of ceteris paribus exercise we would like to approximate.

Given these concerns, a natural choice is to consider the scores of these students in the immediately subsequent course in the sequence. Some care is required in implementing the approach. For instance, it is reasonable to limit the comparison to students who share a common $\kappa = 2$ professor to avoid confounding effects from different $\kappa = 2$ instructors. But even then, one might be concerned about separating the contribution of this common

$\kappa = 2$ professor in the observed score differences across the students. Proposition 2 states that this last point is not an issue as the contribution of the $\kappa = 2$ instructor can be filtered out from the accounting under the injectivity of its production function. Figure 5 captures this intuition graphically (while omitting from the notation $\mathbf{x}_1$ and $\mathbf{x}_2$).

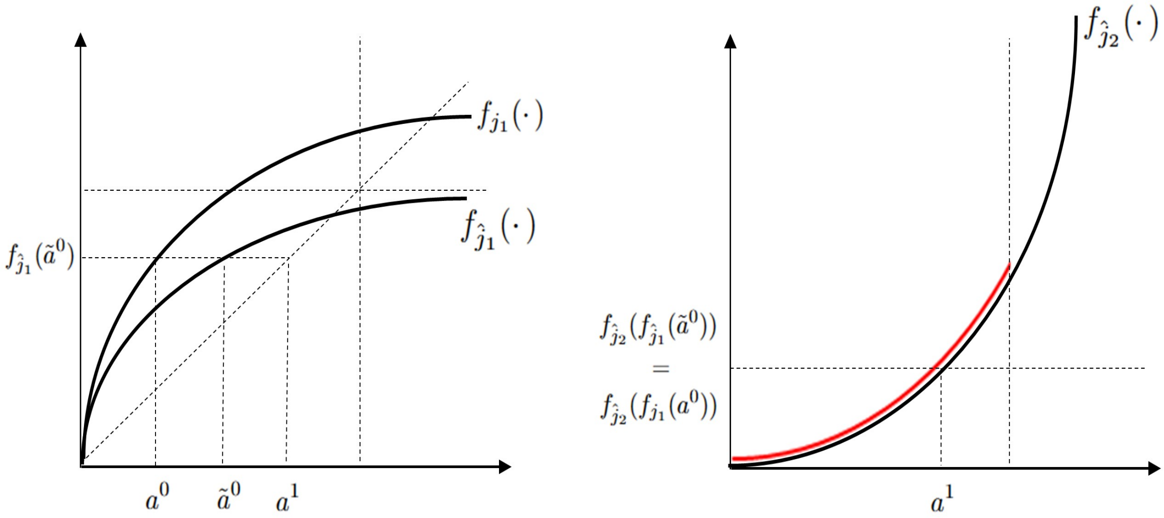


Figure 5: Identification argument for $f_j(a, \mathbf{x})$

Notice that the same arguments in Proposition 7 lend themselves to a partial identification argument in the absence of a type $\tilde{a}_0$ as required by the premise. For example, let's entertain a situation in which $f_{\hat{j}_2}(f_{j_1}(a_0, \mathbf{x}_1), \mathbf{x}_2) > f_{\hat{j}_2}(f_{\hat{j}_1}(\tilde{a}_0, \mathbf{x}_1), \mathbf{x}_2)$ for all types $\tilde{a}_0$ whose validity can be directly observed from the data. The logic of the proof above suggests there is still a lot of information we can extract from this inequality. For instance, assuming $f_{\hat{j}_2}(\cdot, \mathbf{x}_2)$ is monotone increasing for a fixed $\mathbf{x}_2$, we can conclude that $f_{j_1}(a_0, \mathbf{x}_1)$ must exceed the learning return induced by instructor $\hat{j}_1$ under any student $\tilde{a}_0$ in record. In other words, we can construct a lower bound for the unknown $f_{j_1}(a_0, \mathbf{x}_1)$. Similar situations can be treated in an analog way.

We conclude the subsection by highlighting that once the images $f_{j_1}(a_0, \mathbf{x}_1)$ are identified, we can identify the grading policies associated to each instructor by going back to our results on the marginal student associated to each score cutoff for $\kappa = 1$. Proposition 6.1 formally states the latter.

Proposition 3. *The grading policy of instructor j^1 (i.e., β_{j^1}, c_{j^1}) is point identified provided that $f_{j^1}(a_0, \boldsymbol{x}_1)$ is known for some $(a_0, \boldsymbol{x}_1)$.*

Proof. Consider the expressions for $\theta(s_l \mid a_0, \mathbf{x}_1, j^1)$ and $\theta(s_l \mid \tilde{a}_0, \mathbf{x}_1, j^1)$ for two different student types such that $f_{j^1}(a_0, \mathbf{x}_1) \neq f_{j^1}(\tilde{a}_0, \mathbf{x}_1)$

$$\theta(s_l \mid a_0, \mathbf{x}_1, j^1) = \frac{s_l - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1}}{\sigma_\eta^1},$$
$$\theta(s_l \mid \tilde{a}_0, \mathbf{x}_1, j^1) = \frac{s_l - \beta_{j^1} \cdot f_{j^1}(\tilde{a}_0, \mathbf{x}_1) + c_{j^1}}{\sigma_\eta^1}.$$

It is easy to see that given the identification of the images $f_{j_1}(a_0, \mathbf{x}_1)$ and $f_{j_1}(\tilde{a}_0, \mathbf{x}_1)$, the two equations above define a system of equations on the unknowns β_{j^1} and c_{j^1}. The unique solution associated to the system identifies j^1's grading policy. ■

6.2 Identifying the demand for course/sections

The preceding section introduces arguments regarding the identification of the production function model. We now turn our attention into the identification of the underlying primitives within the model for the demand of course/sections. We start the discussion the following assumptions regarding the nature of the demand model error terms.

Assumption 2. *The following assumptions are assumed to be satisfied,*

1. $\{\nu_{i,s,t}\}_{i,s,t}$ *is a collection of mean zero i.i.d. random variables whose distribution, denoted by* $F_\nu(\cdot)$, *is known to the researcher.*
2. *The distribution* $F_\nu(\cdot)$ *is continuous and of full support.*
3. *The utility function* $U_{s,t}(s_j(a_{i,t}, \boldsymbol{x}_{i,j,t}), f_j(a_{i,t-1}, \boldsymbol{x}_{i,j,t}))$ *takes the following functional form* $U_{s,t} = \lambda_{s,t} + \alpha_0 \cdot s_j(a_{i,t}, \boldsymbol{x}_{i,j,t}) + \alpha_1 \cdot f_j(a_{i,t-1}, \boldsymbol{x}_{i,j,t}))$.

The first part of the assumption is a standard independence assumption for the demand model error terms. Assuming a common distribution for $\nu_{i,s,t}$ across all indices, is also a standard assumption that allows us map the demand model objects to the observed instructor market shares via simple conditional choice probabilities. The second and third assumptions are technical and simply allows us to borrow some identification results from the discrete choice literature on linear random utility models.

Given these assumptions, our identification argument can be framed in terms of two main observations. First, we emphasize the implicit assumption that the cost associated with participating in the course enrollment mechanism (i.e., $\phi(\tau)$) remains uniform across all students, thereby rendering it independent of the student index i. Consequently, we can conceptualize $\max_\tau\{\phi(\tau);\ \text{s.t. } s \in \mathcal{C}_t(\tau)\}$ as a professor-period specific fixed cost incurred by a student when expressing its preference for instructor j during academic period t. This characteristic proves useful as it enables us to formulate the course demand problem as a standard discrete choice problem with utilities under alternative-period fixed effects.

Denoting the sum term $\lambda_{s,t} + \max_{\tau}\{\phi(\tau);\text{ s.t. } s \in \mathcal{C}_t(\tau)\}$ by $\Phi_{s,t}$, the ex-ante formulation of the demand model for a student enrolling course κ in period t can be written as follows,

$$\max_{s \in Sect_t^{\kappa}} \left[\Phi_{s,t} + \alpha_0 \cdot s_j(a_{i,t}, \mathbf{x}_{i,j,t}) + \alpha_1 \cdot f_j(a_{i,t-1}, \mathbf{x}_{i,j,t}) + \nu_{i,s,t}\right]$$

Second, we draw attention to the nature of the identification arguments put forth in the preceding section, which establish the identification of both the learning production functions and the grading policies associated to each professor. Consequently, when considering the identification of the course demand primitives, it becomes possible to regard $f_j(a, \mathbf{x})$ and $s_j(a, \mathbf{x})$ as observed quantities. These two observations significantly simplify the identification problem of the demand model, reducing it to the identification of a simple Random Utility Model (RUM). The arguments for the identification of the α_0, α_1 primitives are standard so that we can state the following result without a proof.

Proposition 4. *The parameters α_0, α_1, and $\Phi_{s,t}$ are point identified.*

While the parameters α_0 and α_1 correspond to preference primitives, the term $\Phi_{s,t}$ bundles both the preference fixed effect term and the equilibrium cost of demanding section s. Constructing counterfactual simulations involving student choice will require the disentangling of both quantities. With this goal in mind, consider the following assumptions regarding the nature of the equilibrium cost term,

Assumption 3. *The following assumptions are assumed to be satisfied,*

1. *For all undersupplied sections* $\max\{\phi(\tau)\ s.t.\ s \in \mathcal{C}_t(\tau)\} = 0$.
2. *The fixed effect $\lambda_{s,t}$ can be decomposed in terms of $\lambda_{s,t} = \lambda_j + \boldsymbol{\gamma}' \boldsymbol{z}_{s,t}$ where λ_j where λ_j is a professor fixed effect and $\boldsymbol{z}_{s,t}$ is a vector of observable section characteristics.*

Both assumptions can be explained in intuitive terms. For example notice that at equilibrium, if a section s is undersubscribed any student demanding the section could enter the platform at any period τ and still find a slot in the section. The model predicts that in such cases the student will choose to enroll the section at utility the global max described by $\max_{\tau}\{\phi(\tau)\ s.t.\ \tau \in \mathcal{C}_t(\tau)\}$. The first part of Assumption 3 normalizes the location of utility in terms of such quantity. The second part of the assumption, although restrictive, can be easily explained in compelling economic terms. Intuitively, the assumption restricts the unobserved fixed effects of utility to take one of two forms: (i) u unobserved characteristics at the professor level as opposed to the section level [6], and (ii) unobserved characteristics at the section level that can be mapped into a vector of observed covariates $\mathbf{z}_{s,t}$. Reasonable things to be considered as part of this vector are the

[6]The implication being that such a term would be common to the utility of two sections that share a common instructor

schedule of the course during the day, past student-evaluations for the professor regarding non-learning/scoring attributes, among others.

Assumption 3 allows us to make progress with the identification now of the preference parameters λ_s and $\boldsymbol{\gamma}$. The intuition behind the argument can be described in terms of two steps. First, the first part of Assumption 3 implies that for undersubscribed sections, our identification of $\Phi_{s,t}$ corresponds to identifying the quantity $\lambda_j + \boldsymbol{\gamma}^{'}\mathbf{z}_{s,t}$. It is direct to see that variation of the vector $\mathbf{z}_{s,t}$ within any given undersubscribed professor identifies the vector $\boldsymbol{\gamma}$. Second, given the knowledge of $\boldsymbol{\gamma}$, we can identify the professor fixed effect λ_j for any instructor who is undersubscribed at least one term in the sample. The formal argument is given in what follows.

Proposition 5. *Let j be a Calculus 1 instructor leading two undersubscribed sections, s and $s^{'}$, in two academic periods, t and $t^{'}$. The following results then hold.*

1. *$\boldsymbol{\gamma}_k$ is identified provided that $\boldsymbol{z}_{k,s,t} \neq \boldsymbol{z}_{k,s^{'},t^{'}}$ and $\boldsymbol{z}_{\tilde{k},s,t} \neq \boldsymbol{z}_{\tilde{k},s^{'},t^{'}}$ for all $\tilde{k} \neq k$.*
2. *Given knowledge of $\boldsymbol{\gamma}$, λ_j is identified for any instructor who leading at least one undersubscribed section.*

Proof. Let j be an instructor who leads two undersubscribed sections, s and $s^{'}$, in two academic periods, t and $t^{'}$. Under the first part of Assumption 3 we have that,

$$\Phi_{s,t} = \lambda_j + \boldsymbol{\gamma}^{'}\mathbf{z}_{s,t}$$
$$\Phi_{s^{'},t^{'}} = \lambda_j + \boldsymbol{\gamma}^{'}\mathbf{z}_{s^{'},t^{'}}$$

Under the assumptions of the first part of the proposition, the difference of the two expressions above returns: $\Phi_{s,t} - \Phi_{s^{'},t^{'}} = \boldsymbol{\gamma}_k \cdot (\mathbf{z}_{k,s,t} - \mathbf{z}_{k,s^{'},t^{'}})$. Since we have already argued that $\Phi_{s,t}$ is identified, the latter equation can be solved for $\boldsymbol{\gamma}_k$ proving the first part of the result. The second claim follows trivially from the fact that now for any instructor j undersubscribed for at least one section s we have $\Phi_{s,t} = \lambda_j + \boldsymbol{\gamma}^{'}\mathbf{z}_{s,t}$. Given knowledge of $\boldsymbol{\gamma}$, all terms in the previous equation except for λ_j are identified so that we can solve for the latter. ■

7 Estimation and results

7.1 Parameterizing the model

As discussed above, in principle we could completely estimate the model by implementing a nonparametric estimator based on our identification arguments. While this approach offers certain desirable features, including the ability to refrain from imposing parametric assumptions on key elements such as the learning production functions, practical considerations make the idea of a more restrictive parametric approach attractive. For example, on the side of learning production, our arguments rely on the possibility of matching empirical and theoretical moments for subpopulations of students who share common academic paths. Given the relatively modest class sizes in our context, the observed student count within these subpopulations might be insufficient for empirical moments to closely resemble their theoretical counterparts. Analogous concerns may arise within the demand model.

For this reason, we consider here adopting a fully parametric approach for the estimation exercise. This adjustment not only alleviates data limitation constraints but also affords us the opportunity to specify certain parameters of interest as being common to all instructors which further reduces the data demands of the model. In what follows, we delve into the specifics of these empirical model restrictions and explain how we estimate the resulting model. The final subsection presents the estimates resulting from the approach.

Parameterizing the learning production function

The specification of our empirical model commences with the parameterization of the learning production function associated with each instructor. This is guided by two principal considerations. Firstly, the functional form must exhibit enough flexibility as to accommodate a wide array of learning production shapes. Secondly, the model should be able to capture non-trivial matching effects in the learning production process. Specifically, we aim to capture interactions between instructor characteristics and our measure of student ability.

These two concerns respond to the need to allowing for flexibility at the estimation stage so that the model is capable of capturing the true shape of the learning production functions. For example, consider the additively separable specification common in empirical work, $f_j(a_0, \mathbf{x}) = \delta_j + g(a_0)$. The parameterization would be inadequate for our purposes as it would eliminate, at the modeling stage, the possibility of matching effects in the production of learning. As a second example, the multiplicatively separable parameterization $f_j(a_0, \mathbf{x}) = \delta_j \cdot a_0)$, common in theoretical settings, address the previous concern but implies simplistic positive/negative assortative matching as the only plausible learning ideal scenarios. To address these concerns, our proposal is the following: $f_j(a, \mathbf{x}) = \tilde{\delta}_j^0(\mathbf{x}) + \tilde{\delta}_j^1(\mathbf{x}) \cdot a^{\tilde{\delta}_j^2(\mathbf{x})}$. This formulation accounts for differences in the level of the learning production function, the size of the marginal returns to ability, and the nature of the returns to scale to ability. Additionally, it allows each of these coefficients to vary across instructors both in terms of observed and unobserved attributes.

To be concrete consider partitioning the covariate vector as $\mathbf{x}_{i,j,t} = (\mathbf{x}_{1,i}, \mathbf{x}_{2,j,t})$. The first component encapsulates time-invariant characteristics of students. In our estimation exercise we consider $\mathbf{x}_{1,i} = (sex_i, \{maj_{i,d}\}_{d=1}^{4})$ where sex_i is a male dummy variable and $major_{i,d}$ is a dummy indicating whether student i's major choice is part of department d (i.e., we partition the set of all majors in terms of four major departments, closely following the organizational division within the university) one of four major departments in the university. Meanwhile, $\mathbf{x}_{2,j,t} = (load_{j,t}, ten_{1,j,t}, ten_{2,j,t})$. The variable $load_{j,t}$ is a binary variable specifying whether the total number of sections taught by instructor j in the academic period t exceeds a certain threshold. This allows us to account for either positive returns (learning from teaching multiple sections) or negative returns (potentially due to fatigue) associated with an instructor's teaching load in a given term. Furthermore, $ten_{1,j,t}$ and $ten_{2,j,t}$ are binary variables denoting whether or not the instructor's tenure at period t exceeds certain thresholds. The former, $ten_{1,j,t}$, reflects the number of terms of the course sequence that instructor t has taught by the beginning of period t, while $ten_{2,j,t}$ similarly measures tenure across all courses the instructor has taught in the university.

Given these considerations, we parameterize the production function coefficients in terms of the following,

$$\begin{aligned}
\tilde{\delta}_j^0(\mathbf{x}) &= \delta_j^0 + \boldsymbol{\mu}_0^{'}\mathbf{x}_{2,j,t} + \boldsymbol{\gamma}^{'}\mathbf{x}_{1,i}, \\
\tilde{\delta}_j^1(\mathbf{x}) &= \delta_j^0 + \boldsymbol{\mu}_1^{'}\mathbf{x}_{2,j,t}, \\
\tilde{\delta}_j^2(\mathbf{x}) &= \delta_j^2 + \boldsymbol{\mu}_2^{'}\mathbf{x}_{2,j,t}.
\end{aligned}$$

. Here, $\delta_j^l;\ \ l \in \{0, 1, 2\}$ are production fixed effects capturing unobserved ways in which specific instructors influence production output. In turn the terms $\mu_{.}^{'}\mathbf{x}_{2,j,t}$ capture productivity differences arising from observed heterogeneity reflected in $\mathbf{x}_{2,j,t}$. Importantly, notice the t index in the latter suggesting these variables change over time, thus allowing them to be separately identified from the instructor fixed effects parameters. Finally, $\gamma^{'}\mathbf{x}_{1,i}$ allows for differences in student characteristics other than ability to affect the level of the learning production function.

To complete the description of the learning output empirical model, we must specify the distribution of the error terms in the learning production model. Given that we interpret these as perturbations of the scoring and course dropping equations, a reasonable distributional assumption is $\eta_{i,j,t}^{\kappa} \sim \mathcal{N}(0, \sigma_{\eta}^{\kappa})$ and $\varepsilon_{i,j,t}^{\kappa} \sim \mathcal{N}(0, \sigma_{\varepsilon}^{\kappa})$.

Parameterizing the course/section demand model.

It remains is to specify a concrete distributional form for the error terms $\nu_{i,j,t}$ capturing heterogeneity in taste. We assume these distribute $\nu_{i,s,t} \sim TIEV$. The resulting conditional choice probabilities are of the standard logit form as considered below for a student who demands a section s under instructor j in academic term t.

$$\mathbb{P}(s \mid a_{i,t-1}, \mathbf{x}, t) = \frac{\exp\big(\Phi_{s,t} + \alpha_0 \cdot s_j(a_{i,t-1}, \mathbf{x}_{i,j,t}) + \alpha_1 \cdot f_j(a_{i,t-1}, \mathbf{x}_{i,j,t})\big)}{\sum_{s'} \exp\big(\Phi_{s',t} + \alpha_0 \cdot s_{j'}(a_{i,t-1}, \mathbf{x}_{i,j',t}) + \alpha_1 \cdot f_{j'}(a_{i,t-1}, \mathbf{x}_{i,j',t})\big)}.$$

7.2 Estimation via maximum likelihood

Our parametrization of the model and the distributional assumptions over the error terms suggest a simple estimation approach via Maximum Likelihood (ML). Under this approach we could proceed by maximizing the log of the likelihood associated to our observed data as specified in the following expression,

$$\mathcal{L}(\theta) = \sum_{i=1}^{N} \sum_{t=t_{i,0}}^{T_i} log\Big[\mathbb{P}\Big(j_t \ , \ S_{i,j_{i,t}}^{\kappa_{i,t}} \ \Big| \ j_{i,\tau}, ..., j_{i,t_{i,0}}, a_{i,t_{i,0}}, \mathbf{x}_{i,j_{i,t},t} \ ; \ \theta \Big)\Big].$$

where θ denotes the vector of all parameters associated to each of the courses in the sequence $\mathcal{K}$ considered. In practice however, this approach faces some problems. First, considering the estimation of parameters for all courses simultaneously might pose numerical complications simply due to the number of these parameters. Second, notice that since we don't observe a student's type except for the initial ability $a_{i,t_{i,0}}$ as measured by the entrance exam record, a student's type increases across time. For example, a student's type in its second academic period can be thought as a pair $(a_{i,t_{i,0}}, \mathbf{x}_{i,j_{i,t_{i,0}}}, j_{i,t_{i,0}+1})$, describing the student's initial ability, the instructor the student is paired in its initial period $t-_{i,0}$, and the vector of covariates under which such a match takes place. Even for a modest number of courses in the sequence, the resulting computations required to code the gradient, as required for the implementation of an optimization routine, can be very cumbersome.

Faced with these considerations, we we opt for a sequential ML approach that involves iterating over the different courses in $\mathcal{K}$. At the κ-th iteration of the approach er estimate the primitives θ^{κ} associated to the κ-th course in the sequence. We then use these estimates, in particular those pertaining course κ's production functions, to update each student's ability as implied by the model. The latter estimates then serve as the basis for the $\kappa + 1$-th iteration of the algorithm where they are taken as observed data.

To formally illustrate this approach, let θ^{κ} denote the set of parameters associated with the κ-th course of the sequence $\mathcal{K}$. This includes both primitives of the course/section demand model and primitives of the learning production function model. We begin by constructing the log-likelihood based on the observed data for each student's first two consecutive academic periods upon enrolling in the university. For instance, if student i enrolls in course $\kappa = 1$ for the first time in academic period t, we utilize the data corresponding to academic terms t and $t+1$ for this student. Based on this, the data explained by the model for a student i with $t_{i,0} = t$ is given by: (i) $a_{i,t}$, the student's initial ability, (ii) $j_{i,t}$,$\mathbf{x}_{i,j_{i,t},t}$, the student's instructor match and vector of covariates for the first course enrollment instance, and (iii) $j_{i,t+1}$,$\mathbf{x}_{i,j_{i,t+1},t+1}$, the student's instructor match and vector of covariates for its second academic period. The average loglikelihood function for this data can be expressed in terms of the following,

$$\mathcal{L}(\theta^1, \theta^2) = \sum_{i=1}^{N} \sum_{t=1}^{2} \; log \Big[\mathbb{P}\big(j_{i,t} \mid a_{i,t-1}, \mathbf{x}_{i,j_{i,t},t} \; ; \; \theta^t\big) \cdot \mathbb{P}\big(S^{\kappa_{i,t}}_{i,j_t}, R^{\kappa_{i,t}}_{i,j_t} \mid j_{i,t}, a_{i,t}, \mathbf{x}; \theta^t\big) \Big].$$

where with a slight abuse of notation, we use $t \in \{0,1\}$ to refer to the student i's first and second academic term as opposed to the actual academic period corresponding to these two enrollment instances. The first term inside the logarithm corresponds to the likelihood of observing the student's demand for the section they enroll in during academic term t. The second factor represents the likelihood of the observed course outcomes achieved by the student upon enrolling with a particular instructor in that term. Given that we only consider the first two academic terms of each student, our data contains observations solely for the first two courses in the sequence. Consequently, the log-likelihood function provided above depends exclusively on the parameters associated with these initial two courses: θ^1 and θ^2.

Implementing the sequential ML estimator for θ^1, θ^2 requires addressing a small identification concern. In essence, recall that in our identification results, disentangling the grading policy slopes (i.e., β_j) and the production function images (i.e., $f_j(a_{i,t}, \mathbf{x}_{i,j,t})$) for instructor in $\kappa = 2$ requires information on the student's performance on course $\kappa = 3$. However, the proposed sequential ML approach, the log-likelihood described above does not consider such information. This limitation arises because students can potentially reach course $\kappa = 3$ only in their third academic period, assuming they do not fail the first two courses in the sequence. This issue persists even if we were to estimate all parameters simultaneously, disregarding the sequential ML estimator, as it would still require the decoupling of grading policies and learning outcomes for the last course in the sequence considered.

To address this issue, we reparameterize the model for course $\kappa = 2$ in a way that renders an identified model while keeping θ^1 unchanged. Specifically, let's impose the restriction of $\beta_{j_2} = 1$ and c_{j^2} for all $\kappa = 2$ instructors. This accounts to defining the model for the second period in terms of production functions $\mathring{f}_{j^2}(a, \mathbf{x}) = \beta_{j^2} \cdot f_{j^2}(a, \mathbf{x}) + c_{j^2}$ that pool both learning output and grading policies into a single object. We denote the resulting vector of parameters for $\kappa = 2$ by $\mathring{\theta}^2$. Importantly, notice that the resulting model for student's first two academic terms can be used to identify the parameters θ^1 as our identification of the $\kappa = 1$ parameters doesn't require distinguishing $\kappa = 2$'s grading policies from production function images. We can thus optimize the log-likelihood function under the proposed parameterization and obtain consistent estimates for θ^1, $\hat{\theta}^1$.

Once the latter is achieved, we can transition to the second stage of the sequential ML approach where we estimate the parameters of the second course in the sequence. To accomplish this we use the estimates $\hat{\theta}^1$ for the learning production functions associated to any instructor j^1. These allow us to compute estimates for the implied learning outputs associated to each student's first academic period match. For example, after enrolling professor j^1's section for $\kappa = 1$, a student with an initial ability measurement of $a_{i,0}$ ends up with a new ability given by $a_{i,1} = f_{j^1}(a_{i,0}, \mathbf{x}_{i,j^1,1}\,; \hat{\theta}^1)$ where the notation makes it clear that we use the first step estimates in order to construct these ability estimates. At this point we can treat $a_{i,1}$, the ability estimates from the previous stage as observed data and use them to estimate θ^2 in the current stage of the algorithm. The process follows the same steps as before: (i) constructing the log-likelihood using data from two consecutive periods for all students upon their enrollment in a course $\kappa = 2$, and (ii) reparameterizing the model for θ^3 to account for the lack of identification of grading policies/ production functions for $\kappa = 3$.

7.3 Estimation results

This section introduces the main findings resulting from the estimates of the model. The discussion is organized in terms of three main components: (i) estimates for the average learning production function, (ii) estimates for the distribution of learning outputs across professors, and (iii) estimates for the demand model primitives. In all three cases, emphasis is placed on the implications of the estimates over the observed student-professor assignment and the possibility of improving such assignment via counterfactual policies explored in subsequent sections.

Average learning production function estimates

Let's begin by documenting the average learning production function where the average is considered across the population of Calculus 1 instructors. This provides a compact way of summarizing the average relationship between learning outcomes and students' inputs to learning. Figure 6 depicts these average learning production outcomes as a function of a student's ability. Each panel conditions on a different value for the vector of covariates $\mathbf{x}$. For example, the second panel describes the subpopulation of male students majoring in STEM fields who enroll in Calculus 1 under an instructor characterized by am above average teaching load and an above average course-specific/general tenure. The x-axis of each panel displays a student's ability level. In turn the blue curve's value at any given ability level represents the average of the images for each instructor's learning production function according to the estimates (i.e.,

$\hat{f}_j(a, \mathbf{x})$). The shaded area corresponds to the associated 95% confidence interval.

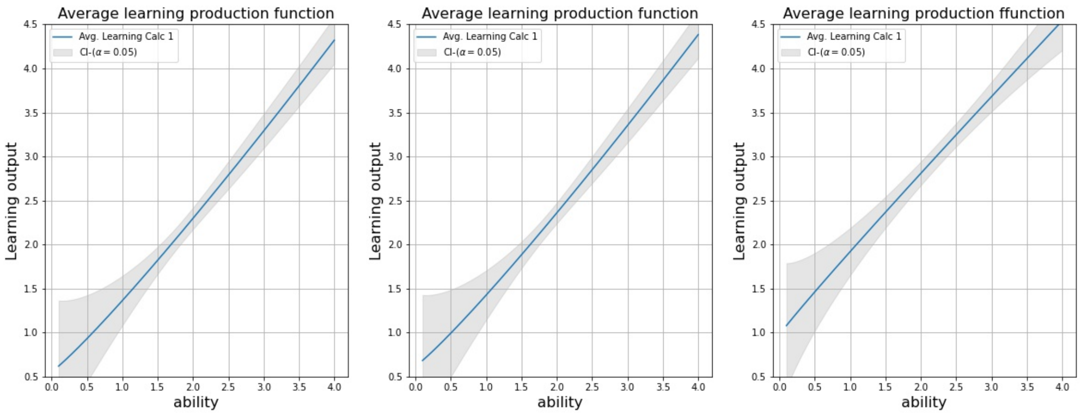


Figure 6: Average learning production function

Notes: Panel 1 considers male students in the excluded major category under high teaching load, high course tenure, and high general tenure. Panel 2 coincides with the latter except that it focuses on STEM students. Panel 3 pairs female students in the excluded major category with professors having low teaching load, low course-specific tenure, and low general tenure.

Two key observations emerge from these plots. First, each panel illustrates significant disparities in learning outcomes among students of varying ability levels. To put on numbers on this claim, consider in the first panel a student of ability $a_i = 1.0$, which is at the lower end of the ability spectrum. When randomly assigned to one of the Calculus 1 instructors in the data, such a student can anticipate an average learning outcome of approximately 1.3 GPA points under the reference grading policy, or equivalently, a D grade. In contrast, a student of an ability of $a_i = 4.0$, positioned at the top of the distribution, achieves a learning outcome of approximately 3.8, equivalent to a B+ grade. More generally, the average learning production function is increasing relative to a student's ability level, and spans a wide range of distinct learning outputs. This is robust to changes in the vector of covariates as suggested by looking at the remaining panels.

Second, factors beyond a student's ability type influence the shape of the average production function. For instance, while the first and second panels exhibit slightly convex relationships, the last panel portrays a scenario characterized by diminishing returns to a student's ability, as evident from the modest concavity of the average production function. In addition, the overall height of the production function varies across these panels. Consider for example a student with an ability of $a_i = 2.0$. In the first panel, this student achieves an approximate learning output of 2.0 GPA points under the reference professor, while in the second panel, its average learning output increases to approximately 2.5. This students output is even higher under the third conditioning set, reaching approximately

3.5 GPA points under the reference grading policy. In the appendix I show that these observations correspond to the estimates of the slope parameters for the covariate variables for which point estimates and standard errors are reported.

The distribution of learning outcomes across instructors

Gains from our reassignment counterfactual exercises depend on the existence of teaching ability differences across instructors. To understand these differences we need to look not at the average in the distribution of learning outcomes, but at the dispersion of the distribution. The left panel of Figure 7 shows this by plotting various percentiles in the distribution of learning outcomes. As before, the x-axis corresponds to the student's ability level. For any given ability, the images of the curves being depicted represent percentiles in the distribution of learning outputs for a student of such ability. While the Figure corresponds to a fixed vector of covariates, the Appendix section shows the same conclusions are reached after conditioning on other covariates.

Consistent with the stylized facts reported before, non-negligible variation exists in the learning outcomes a student can expect when paired with a Calculus 1 instructor. For instance, a student with an ability level of $a_i = 2.0$ can expect an average learning outcome of 2.1 GPA points when randomly paired with a Calculus 1 instructor. However, the range of possible learning outcomes extends from 2.0 GPA points at the 10th percentile to 2.5 GPA points at the 90th percentile. The latter difference is large, being equivalent to transitioning from a D to a C grade under the reference instructor's grading. The dispersion in learning outcomes becomes even more pronounced when considering students at the higher end of the ability distribution. For example, the same percentiles for a student with an ability of $a_i = 3.0$ correspond to a two-letter grade jump, ranging from 2.7 GPA points to 3.5 GPA points, or equivalently from C+ to B+ in terms of letter scores.

The second panel of Figure 7 complements the latter by displaying the distribution of scoring outcomes that our randomly assigned student can anticipate. This provides an intuitive way of incorporating differences in instructors' grading policies into the presentation of our estimates. To start, notice that the average score output across instructors closely mirrors the average learning output curve in the first panel. This is consistent with the grading policy estimates, reported in the appendix, which suggest the average instructor's grading policy coincides with that of the reference professor. However, it's worth highlighting that the spread in the scoring distribution is higher than that of the distribution for learning outcomes. This holds true for all student ability levels and remains robust across changes in the conditioning vector of covariates, as illustrated in the appendix. For instance, for the student of ability $a_{i,0} = 2.0$ considered before, a move from the 10th percentile to the 90th percentile in the distribution of scores corresponds to a jump from approximately 1.6 to 2.6 GPA points.

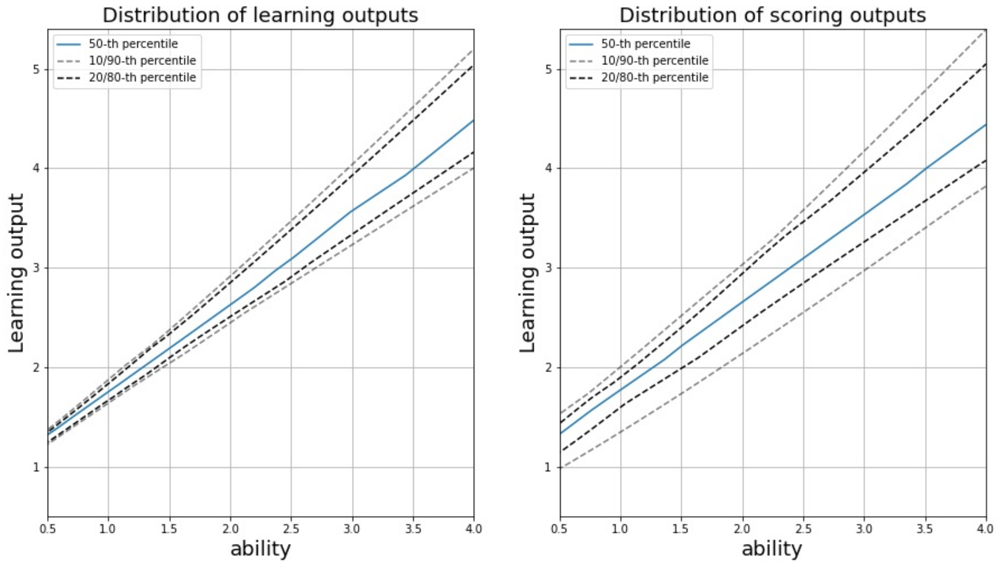


Figure 7: Distribution of learning outcomes

Notes: Both panels correspond to a female student in a STEM major, under above average teaching load, course tenure, and general tenure instructor.

Collectively, the estimates in Figure 7 indicate a potential source of tension for students concerning their preferences for both learning and scoring outcomes. In other words, if instructors associated with high-learning outcomes differ from those with high-scoring outcomes, students will encounter a trade-off between learning and scoring when selecting sections of Calculus 1. The ultimate choice a student makes will depend on the weight its preferences place on each of these aspects. Before delving into the preference side of this issue, let's analyze this trade-off by simply documenting the extent to which the optimal teaching and scoring instructors for a student differ, as well as the average learning/scoring magnitudes of these differences.

For instance, consider computing the score and learning outcomes for each course enrollment instance for Calculus 1 in our sample. We can then determine for each instance the number of instructors associated with a higher scoring output than that of the student's optimal professor in terms of learning output. Table 5 records this information. Each column in the table corresponds to a different integer value representing the number of professors who would improve a student's score, relative to the best instructor for the student in terms of the induced learning. The data in the k-th column for a given row is interpreted as the proportion of students within the subpopulation represented by the row who have k score-improving instructors relative to the learning optimal professor. To highlight variations in these score-improving opportunities across students of different abilities, the information is shown for different segments of the student ability distribution.

For the population of students as a whole, a large fraction of all Calculus 1 course-enrollment instances are such that the best instructors in terms of learning and scoring don't coincide. For instance, across all the course enrollment instances, 85% are associated to at least one score-improving instructor. Significant differences result from considering students of different ability levels, with the highest number of score-improving professors showing up among students at the top of the ability distribution. To some extent, the latter reflects the higher dispersion in the scoring and learning distributions according to the estimates.

Table 5: # of score-improving professors relative to the learning optimal professor.

	# of score-improving professors				
	1	2	3	4	≥5
All students	11.00	23.18	9.78	14.57	28.01
Ability 0% - 25%	16.08	22.95	11.42	16.49	8.21
Ability 25% - 50%	8.06	30.00	11.59	17.96	14.24
Ability 50% - 75%	9.78	23.26	8.39	12.52	38.49
Ability 75% - 100%	10.11	16.42	7.69	11.24	51.34

Notes: The k-th column cell for a given row reports the fraction of students in the subpopulation described by the row with exactly k score-improving instructors relative to the learning optimal one.

It is also of value to think about the magnitude of the learning and scoring differences between the learning-optimal and scoring-optimal instructors. Intuitively, the larger the score gap corresponding to the deviation the higher is the temptation to deviate [7] Table 6 considers these two differences. The first column measures the difference in learning outcome between the learning-optimal and the scoring-optimal professor, which here is denoted as the scoring gap. The second column constructs the difference between the scoring outcomes of the scoring-optimal and the learning-optimal instructors, denoted as the scoring gap. Here, the results are revealing, as in all cases the scoring gap substantially exceeds the learning gap. To put this in context, let's entertain a Calculus 1 student of an average average level. While opting for the best instructor in terms of learning involves a premium of approximately 0.32 GPA learning points, opting instead for the best scoring professor involves a premium of 0.91 in terms of scoring.

[7]This depends on the student's placing positive weight over the scoring outcomes. I show this is the case in the following subsection.

Table 6: Learning/Scoring gap between the learning and scoring optimal professors.

	Outcome gaps	
	Learning gap	Score gap
New students	0.32	0.91
Ability 0% - 25%	0.17	0.62
Ability 25% - 50%	0.29	0.76
Ability 50% - 75%	0.35	1.01
Ability 75% - 100%	0.42	1.16
Old students	0.17	0.76
Ability 0% - 25%	0.10	0.43
Ability 25% - 50%	0.15	0.62
Ability 50% - 75%	0.20	0.98
Ability 75% - 100%	0.29	1.28

Notes: The learning gap is defined as the average difference between the learning output a student obtains under the learning optimal instructor and the scoring optimal instructor for the academic term in which the course enrollment instance takes place. The scoring gap is the average difference between the score output a student obtains under the scoring optimal instructor and the learning optimal instructor for the academic period of the course enrollment instance considered.

The intuition above can also be framed in terms of the fraction of students who end sup being paired with the learning-optimal instructor in the data. For new students, this proportion is approximately 12.17%. Notably, this aligns with the fact that first-time students are assigned to sections within Calculus 1 randomly, and the average number of sections in the term when these students enroll in a course is 8.7. This corresponds to an 11.46% probability of being randomly matched with the learning-optimal instructor, very close to our estimate. Regarding students repeating the course, the fraction resulting from the estimates is approximately 16.20%. While a fraction above the case of randomly matched students, still very low.

Course/section demand parameters

The key parameters of interest in our demand model are α_0, representing the marginal utility of expected scores, and α_1, indicating the marginal utility of expected learning outcomes. Table 7 reports the point estimates and standard errors for both parameters under Calculus 1 and Calculus 2 courses. For Calculus 1, both α_0 and α_1 exhibit positive values, signifying that students assign positive weight to both learning and scores when making course/section selections. Notably, the magnitude associated with α_0 is slightly above that of α_1 suggesting students place a higher weight on the students they expect to obtain above their pure learning outcomes.

Table 7: Estimation results - Course/section demand parameters

	Calculus 1	
Parameter	Estimate	Std. Error
α_0	1.19	0.05
α_1	1.17	0.10

We also want to highlight that these estimates indicate a substantial portion of the variation in demand decisions can be attributed to both the scoring and learning dimensions. For instance, when considering the other utility component, $\Phi_{s,t}$, the average value for these quantities among observed (s,t) pairs is -0.35, with a standard deviation of 1.78. These values are of the same order of magnitude to the score and learning outputs for most students in the sample. An important corresponds to low-ability students for whom learning and scoring outputs are small in terms of their magnitude. This suggests that their decisions are relatively more influenced by the fixed effects terms $\Phi_{s,t}$, a possibility we explore when interpreting out counterfactual exercises.

8 Counterfactual - Dictatorial assignment

Describing the counterfactual policy

Consider the problem of matching students and instructors under a dictatorial approach. This resembles the task of a university administration that seeks to pair students with instructors for a section of a course without taking into account students' preferences regarding their instructors. The administration's primary goal is to maximize the average student learning that results from the chosen assignment, as implied by the estimates of the learning production function. I am focusing on a myopic assignment problem in which matches are made on a period-by-period basis. This should be contrasted with the fully dynamic version where the assignment decision in one period considers its impact on the future pool of students seeking to enroll in the course (i.e., the set of students who fail or drop a course changes with different matches). The former approach significantly simplifies the computational challenges of the problem and can be understood as providing a lower bound for the improvements a planner can achieve through reassignments.

To be specific, let's consider the university's problem for the first course in the sequence, Calculus 1, during academic period t. Denoting the pool of students seeking to enroll in a section of the course as $\mathcal{I}_t$, the university's objective is to select $\mu_{i,s} \in 0, 1$ for every $i \in \mathcal{I}_t$. In the latter, $\mu i, s = 1$ is a dummy variable indicating that student i is assigned to section s according to the administration's choice of match. When making these assignments, the university is bound by the exogenous capacity constraint for each section s, ensuring that it is not exceeded. To account for potential distributional concerns regarding the planner's objective, I examine various versions of the problem, differing in the weight, $\omega(a)$, assigned by the planner to each student's ability type. Denoting C_s as the capacity constraint for section s and j_s as the instructor associated with such a section, we can formulate the problem for period t as in the following,

$$\max_{\{\mu_{i,s}\}_{i,s}} \sum_{i \in \mathcal{I}} \sum_{s \in Sect_t} \omega(a_{i,t-1}) \cdot f_{j_s}(a_{i,t-1}, \mathbf{x}_{i,j_s,t}) \cdot \mu_{i,s},$$

subject to the constraints:

$$\sum_i \mu_{i,s} \leq C_s; \quad \forall\, s \in Sect_t,$$
$$\sum_s \mu_{i,s} \leq 1; \quad \forall\, i \in \mathcal{I}_t.$$

For the period t problem, the number of active sections (i.e., $Sect_t$) and the set of instructors leading these sections are exogenously given. In particular, in the simulation exercise, these coincide with the observed sections/instructor pairs for the corresponding academic period t. Given this, I can compute the learning outputs $f_{j_s}(a_{i,t-1}, \mathbf{x}_{i,j_s,t})$ for each potential student-professor match that could result from the planner's assignment choice. It is easy to see that this reduces the planner's problem to a simple assignment problem that can be solved using standard linear programming techniques[8].

After the planner chooses a specific assignment $\mu_{i,s_{i,s}}$, learning takes place according to the estimated production functions $\hat{f}_j$. In addition, students' score outcomes and retirement decisions are simulated by drawing from the estimated distributions for the scoring and dropping equation's error terms. These simulated outcomes determine the total number of students who need to re-enroll in the course in the subsequent term (i.e., due to either failing or dropping a course), which together with the first-time students observed in the data for period $t + 1$ determines $\mathcal{I}_{t+1}$. The planner then repeats the process by solving the assignment problem for students in period $t+1$ as described above.

Simulating the counterfactual policy

[8]The existence of an integer solution is guaranteed for assignment problems of this nature

Figure 8 shows the outputs resulting from the reassignment exercise. Three distinct reassignment simulations are considered, each characterized by a different weight function. In particular, I consider a weighting function $\omega(a) = \exp^{-\delta \cdot a}$, where $\delta \in 0.0, 0.5, 2.0$. The first value in the set corresponds to a uniform weight, placing the same value for all student ability types in the reassignment exercise. In contrast, the latter two assign more weight to students with lower initial abilities, capturing a university's potential concern for improving learning outcomes for disadvantaged students. The first panel illustrates the average learning output for each student ability level. While the average under the observed assignment is represented by a dashed line, the average under each counterfactual reassignment corresponds to a solid line. In addition, the second panel depicts the average learning difference between the counterfactual and the observed assignments, providing a clearer perspective on how various ability types experience gains or losses after the implementation of each policy.

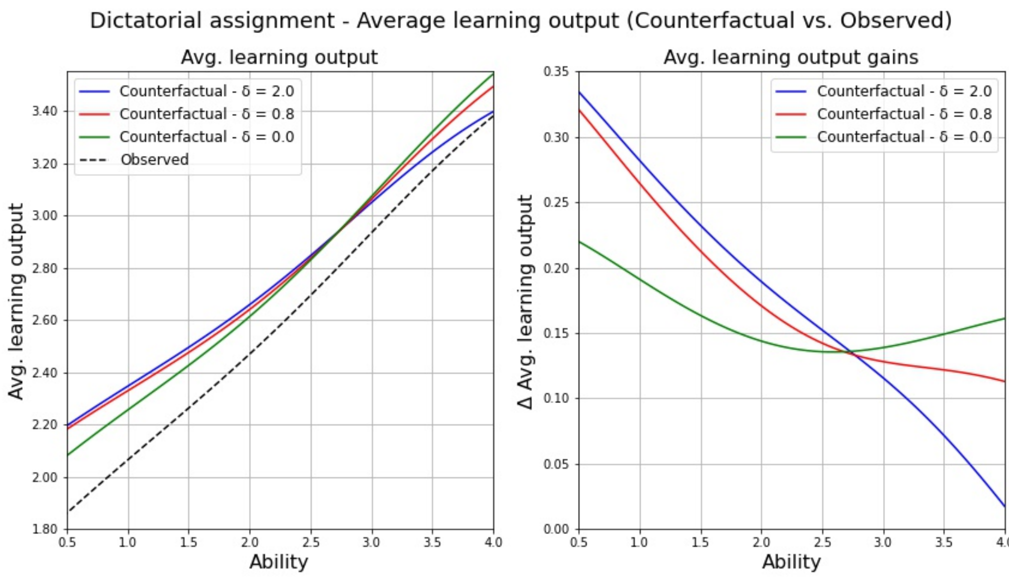


Figure 8: Dictatorial Counterfactual - Avg. learning output

Two observations can be highlighted from the latter exercise. First, it is clear that reassignments improve average learning outcomes, with the magnitude of these gains differing across students with varying initial abilities. For instance, consider the green curve corresponding to the uniform weight reassignment exercise. For students with ability levels described by $a_i \leq 1.0$, the gains correspond to an increase of more than 0.20 GPA points. Students with an ability level of 2.5 exhibit the smallest average gain, approximately 0.14 GPA points. Naturally, as I consider exercises that assign higher weight to students with lower abilities, reassignment gains rise for the latter but diminish for high-ability students. For instance, under the extreme case of $\delta = 2.0$, which assigns very small weight to students with abilities exceeding 2.0, a student with an ability level of 4.0 witnesses no gains relative to the baseline, whereas the average learning return for a student with low ability, such as 1.0, increases to 0.27 GPA points.

Second, I emphasize that while learning gains vary across different student ability types, these gains remain positive across the entire spectrum of ability. At first, this may appear surprising, as it suggests that the university can improve average learning outcomes for all student types without facing the typical trade-offs associated with prioritizing the learning of one subgroup of students at the expense of another. However, the result becomes more intuitive when considering two key factors. First, under the observed assignment, a large fraction of students enroll in Calculus 1 as first-time students. Consequently, they don't participate in the first-come-first-served mechanism and are instead randomly assigned to sections of Calculus 1. Random assignments are prone to generating many Pareto suboptimal pairings, allowing the planner to exchange slots between two students, i and i', in a way that enhances the learning output of both. Second, even among reassignments involving students who are repeating Calculus 1 (i.e., approximately a third of the students in each academic term), the oversupply of instructors/slots, relative to the total number of students demanding sections, implies there is a significant amount of slackness in the capacity constraints of many sections under the observed assignment. This creates many opportunities in which the planner can modify the assignment of a given student without the need to take a slot away from another student in order to meet the capacity constraint.

In trying to understand the outcomes of the exercise, it is useful to examine how the professor assigned to each student-ability type changes during the reassignment. To illustrate this, consider Figure 9, which shows the conditional ability distribution of students matched with each instructor under the uniform weight reassignment. In this figure, each row corresponds to a specific instructor, while the columns correspond to quartiles in the distribution of student ability. A cell associated to a given row/column is interpreted as the proportion of all students matched with the instructor represented by the row who upon enrollment have an ability level belonging to the quartile specified by the column. I have ordered the instructors lexicographically, considering the percentage of students they enroll in each of the four quartiles (i.e., order first in terms of the percent of students in Q1, then Q2, ...).

Notably, under the counterfactual reassignment, there is a discernible specialization in terms of the professors assigned to students with different ability levels. This is evident from the diagonal structure observed in the heatmap of the first panel. This should be contrasted with the right-hand panel showing the conditional distributions under the observed assignment. In the latter, each instructor teaches a substantial number of students in each ability region, although the distributions are not exactly uniform, as the non first-time students can choose the section they enroll in.

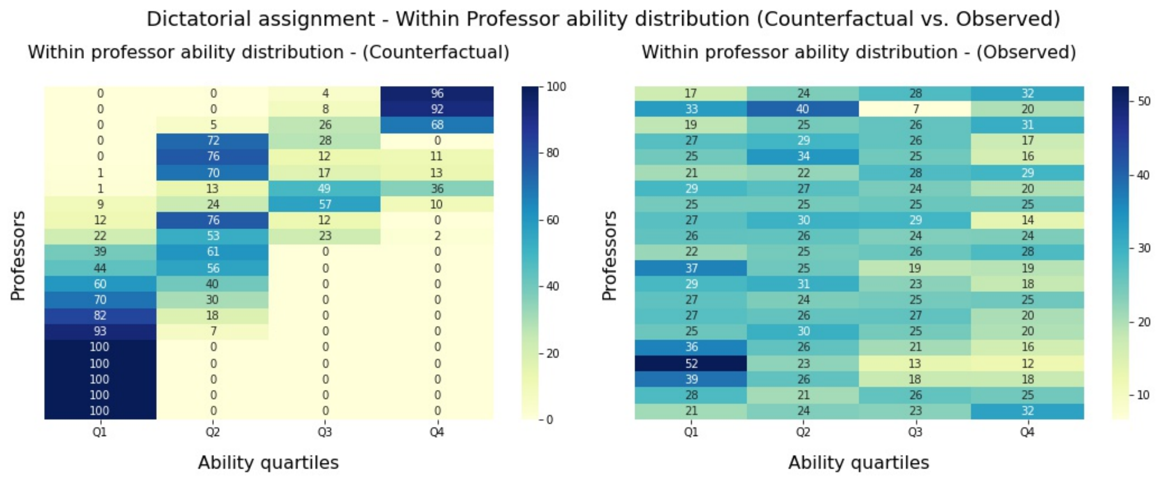


Figure 9: Dictatorial counterfactual - Within professor ability distribution

I reiterate that such specialization does not entail the conventional distributional concerns of aggregating students based on their ability within courses. In other words, the rationale behind the reassignment does not involve the dean prioritizing high learning gains for high-ability students at the expense of low-ability students. As previously discussed, learning gains for low-ability students are actually larger or equal than those observed for high-ability students even under the uniform weight exercise.

Although the reassignment is constructed with average learning outcomes as the objective, the gains are reflected in other related variables of interest to the university's administration. For instance, Figures 10 and 11 depict the changes in the rate at which students withdraw from sections of Calculus 1 and the average number of attempts required for the successful completion of Calculus 1, respectively.

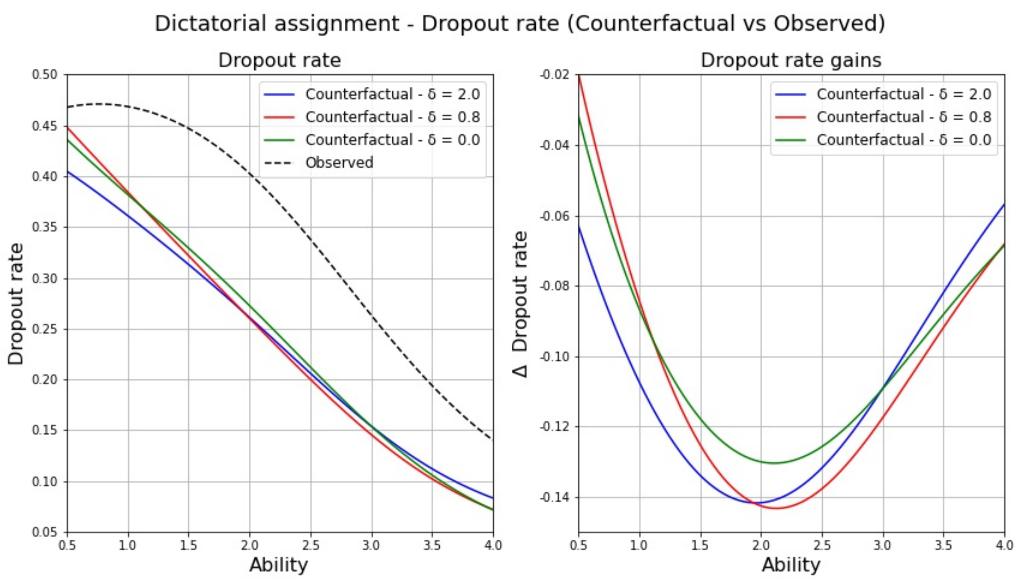


Figure 10: Dictatorial Counterfactual - Course dropping rate

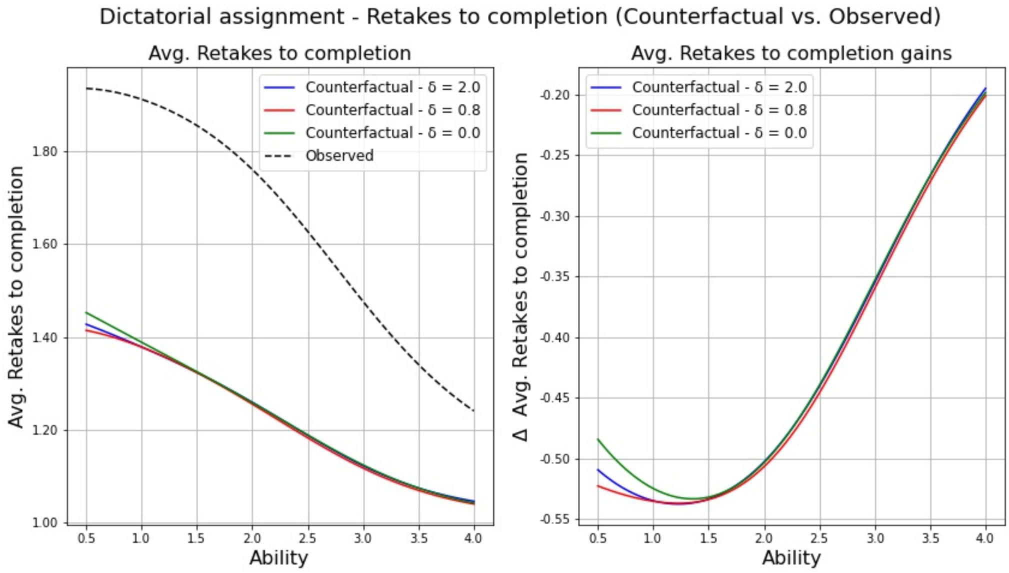


Figure 11: Dictatorial Counterfactual - Attempts to completion

When examining the rate at which students opt to withdraw from their Calculus 1 sections, the reassignment of students yields significant improvements. This is particularly the case for students positioned within the mid-range of the ability distribution. As these are typically the students at the borderline between the course dropping and not dropping decisions, they are also more likely to change their dropping decision after experiencing the boost in their learning outcomes from the reassignment. For instance focusing on the uniform weight reassignment exercise, a student with an ability level of 2.0 drops the Calculus 1 section at a rate of 0.40 under the observed assignment assignment. This rate decreases to 0.27 under the uniform weight dictatorial reassignment, representing a reduction of 0.13 in the odds of dropping.

Turning to the second plot, which illustrates the number of attempts a student requires to achieve a passing score in Calculus 1, we once again observe notable gains resulting from the counterfactual reassignment. For instance, let us revisit the example of a student with an ability level of 2.0, who under the baseline requires an average of 1.70 attempts to successfully complete Calculus 1. Instead, under the counterfactual reassignment, the same student accomplishes this in 1.3 attempts. Substantial gains extend to various ability levels, as indicated by the right panel, with the smallest improvements observed among high-ability students who are less inclined to retake the course to begin with since they perform better. Nevertheless, it is worth noting that even for such students the reassignment leads to a significant drop in the number of attempts to completion. For example, a jump from an average of 1.21 retakes to slightly above 1.0 retakes for students of a top ability 4.0.

9 Conclusion

The learning outcomes resulting from the way in which students and professors are assigned to classrooms depend crucially on the existence and shape of the matching effects in the learning technology. Consequently, the measurement of these complementarities and their role in shaping efficient learning matches becomes a primary concern. This is particularly important in higher-education settings where, for institutional reasons, student-professor assignments are constructed via course-enrollment mechanisms. The learning consequences, while poorly understood at a formal level, can be suspected of generating inefficient assignments due to their reliance on student choice.

In this paper I have studied the measurement of this matching effects and have used the measurements to evaluate the extent to which commonly used course-enrollment mechanisms can produce high-learning assignments. To do this, I construct a structural model characterizing learning outcomes and course/section demand decisions under assumptions common to many post-secondary institutions. My work makes two conceptual contributions to the literature.

First, I provide arguments for the identification of learning technologies in higher-education environments. In particular in situations where instructors differ in their grading policies which invalidates inference of learning production functions by means of the within-professor variation in the scores. Second, I propose a new channel through which grading policies can have real learning consequences by affecting students' demands within choice-based course-enrollment mechanisms. These ideas are illustrated by means of an empirical exercise showing how, for a concrete post-secondary setting, learning outcomes can be improved by means of reassignments that fully exploit the complementarities in the learning production functions. While the empirical results speak to random assignment and first-come-first-served mechanisms, two frequently used mechanisms in practice, the approach is general and can be applied to the study of complementarities in other higher-education environments.

Two unexplored research avenues seem of particular relevance. First, the empirical exercise considered here addresses mechanisms involving random assignment or first-come-first-served schemes. However, other mechanisms are also commonly used in higher-education settings. For example, some universities use lotteries or point systems to allocate sections within a course, while others employ priority systems that allow some students to choose first, aiming to influence the allocation. Exploring the learning consequences of these alternative mechanisms and understanding how they depend on demand patterns remains of interest. In particular, future work could aim to provide richer data descriptions of students' participation in these mechanisms to construct counterfactual comparisons of these different policies. Second, as mentioned before, learning is not the only criterion used to evaluate course-enrollment mechanisms. University administrators very likely have concerns about the distributional and fairness consequences of the mechanisms being used. Additionally, an extensive literature already exists that seeks to compare course-enrollment mechanisms in terms of preference-based notions. An interesting research avenue would involve finding ways to incorporate these concerns into the empirical comparison of these different mechanisms

References

Ahn, T., Arcidiacono, P., Hopson, A., & Thomas, J. R. (2019). *Equilibrium grade inflation with implications for female interest in stem majors* (Tech. Rep.). National Bureau of Economic Research.

Ahn, T., Aucejo, E., James, J., et al. (2020). *The importance of matching effects for labor productivity: Evidence from teacher-student interactions* (Tech. Rep.). Working Paper, California Polytechnic State University.

Aucejo, E. M., Coate, P., Fruehwirth, J., Kelly, S., & Mozenter, Z. (2018). Teacher effectiveness and classroom composition.

Babcock, P. (2010). Real costs of nominal grade inflation? new evidence from student course evaluations. *Economic inquiry*, *48*(4), 983–996.

Budish, E., & Cantillon, E. (2012). The multi-unit assignment problem: Theory and evidence from course allocation at harvard. *American Economic Review*, *102*(5), 2237–2271.

Butcher, K. F., McEwan, P. J., & Weerapana, A. (2014). The effects of an anti-grade inflation policy at wellesley college. *Journal of Economic Perspectives*, *28*(3), 189–204.

Carrell, S. E., & West, J. E. (2010). Does professor quality matter? evidence from random assignment of students to professors. *Journal of Political Economy*, *118*(3), 409–432.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, *104*(9), 2593–2632.

Diebold, F., Aziz, H., Bichler, M., Matthes, F., & Schneider, A. (2014). Course allocation via stable matching. *Business & Information Systems Engineering*, *6*, 97–110.

Figlio, D. N., & Lucas, M. E. (2004). Do high grading standards affect student performance? *Journal of Public Economics*, *88*(9-10), 1815–1834.

Gershenson, S., Holt, S., & Tyner, A. (2022). Making the grade: The effect of teacher grading standards on student outcomes.

Gilraine, M., Gu, J., & McMillan, R. (2020). *A new method for estimating teacher value-added* (Tech. Rep.). National Bureau of Economic Research.

Gilraine, M., & Pope, N. G. (2021). *Making teaching last: Long-run value-added* (Tech. Rep.). National Bureau of Economic Research.

Graham, B. S., Ridder, G., Thiemann, P., & Zamarro, G. (2022). Teacher-to-classroom assignment and student achievement. *Journal of Business & Economic Statistics*, 1–27.

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review*, *61*(2), 280–288.

Hanushek, E. (2009). *Creating a new teaching profession.* ERIC.

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (Tech. Rep.). National Bureau of Economic Research.

Krishna, A., & Ünver, M. U. (2008). Research note—improving the efficiency of course bidding at business schools: Field and laboratory studies. *Marketing Science*, *27*(2), 262–282.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American economic review*, *94*(2), 247–252.

Sönmez, T., & Ünver, M. U. (2010). Course bidding at business schools. *International Economic Review*, *51*(1), 99–123.

10 Appendix

10.1 Identification under course dropping

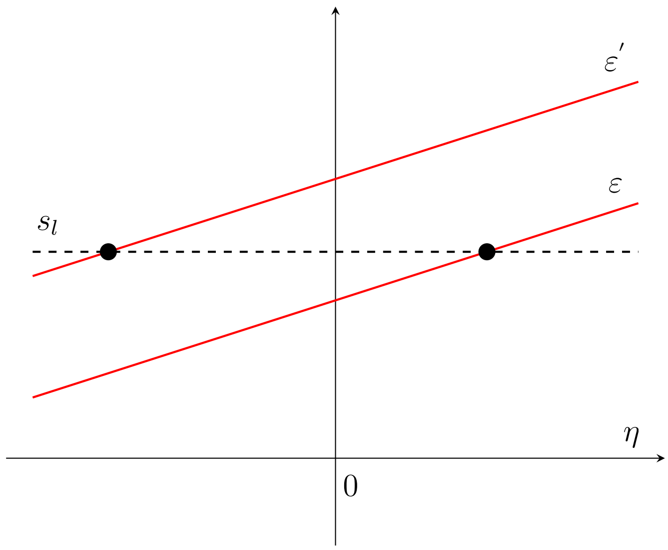
Thus far we have only provided arguments for the identification of our empirical model in a setting where students are unable to drop sections previously enrolled. In situations in which a non-negligible number of students choose to drop a section, ignoring this might be problematic. Intuitively, the issue arises from the fact that all our identification results are based on our observations of the group of students who achieve scores above a certain threshold, denoted as s_l. Yet, when students have the option to drop a course, the researcher can only observe the fraction of students who score above s_l conditional upon not dropping the course.

Failure to account for this distinction can lead to upward bias in our estimators, where estimates suggest that the learning returns of professors are higher than they truly are. Moreover, depending on the relationship between grading policies and professor productivity, this bias can result in erroneous conclusions regarding disparities in learning returns across instructors. For instance, if students are more likely to drop a course under the instruction of subpar professors, we run the risk of underestimating the gap in instructional quality between a high return and a low return instructor.

We now try to extend the arguments in the previous subsection in a way that accommodates for the truncation issue resulting when students can drop a course. Following the same arguments used before, consider the mass of students of initial type $a_{i,0} = a_0$ who in their first academic term obtain a score of at least s_l after enrolling instructor j^1's section under the covariate vector $\mathbf{x}_1$. Following the discussion above, we also condition on the subgroup of students who choose not to drop j^1's section as otherwise we wouldn't observe a score record for the student. Under our model, the latter conditional probability can be written as follows,

$$
\begin{aligned}
&\mathbb{P}\big(S^1_{i,j^1} \geq s_l \mid a_0,\ \mathbf{x}_1,\ j^1,\ R^1_{i,j^1} = 0\big) \\
&= \int_{\varepsilon^1} \int_{\eta^1} \mathbf{1}\big\{R^1_{i,j^1} = 0\big\}\mathbf{1}\big\{S_{i,j^1} \geq s_l\big\}\ f_\eta(\eta^1) f_\varepsilon(\varepsilon^1)\, d\eta^1\, d\varepsilon^1 \\
&= \int_{\varepsilon^1} \mathbf{1}\left\{\varepsilon^1 \geq \frac{s_{l^*} - \beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}\right\}\left[1 - F_\eta\left[\frac{s_l - \beta_{j^1}\cdot f_{j^1}(a^0, \mathbf{x}_1) - c_{j^1} - \sigma^1_\varepsilon \cdot \varepsilon^1}{\sigma^1_\eta}\right]\right] f_\varepsilon(\varepsilon^1) d\varepsilon^1.
\end{aligned} \tag{1}
$$

This expression closely resembles the one derived in our previous identification when trying to identify the marginal student for score s_l. It however differs in two crucial aspects that complicate this interpretation. First, students (within the conditioning set) now differ in two unobserved ways: their η^1 and ε^1 draws. Given that conditional on not dropping the course, ε^1 has an impact on the final score received by a student, this means we now should think not of a single marginal student for score s_l but of a marginal student for s_l under each potential draw of ε^1. Our observation of the mass of student scoring above s_l now corresponds to adding up the mass of students who score above s_l across all this different sub populations based on ε^1. Figure ..., the analog to Figure ... in the main subsection illustrates this by showing how a change in ε^1 shifts the linear function defining the marginal student. Second, the integral on the right hand side must now reflect the fact that we only consider students who choose not to drop the course. This is captured by the indicator term inside the integral $\mathbf{1}\{R^1_{i,j^1} = 0\}$. In terms of the graph below, this amounts to only counting the students who score above s_l for some of the ε^1 sub populations, namely, those who choose not to drop the course.



In practical terms, the implication is that we cannot directly invert the previous equation to learn about σ^1_η and $f_{\hat{j}^1}(a_0, \mathbf{x}_1)$ as considered in our previous arguments. Some additional work is required in order to accomplish this. Nevertheless, what we can do is to identify the marginal student, not in terms of obtaining a score s_l, but in terms of choosing to drop the course. To see this, consider an expression for the mass of students who choose not to drop j^1's section under our conditioning set,

$$\begin{aligned}\mathbb{P}\big(R^1_{i,j^1} = 0, \mid a_0,\ \mathbf{x}_1,\ j^1\big) &= \int_\varepsilon \mathbf{1}\left\{\varepsilon \geq \frac{s_{l^*} - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}\right\} f_\varepsilon(\varepsilon) d\varepsilon \\ &= 1 - F_\varepsilon\left[\frac{s_{l^*} - \beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}\right]\end{aligned}$$

Inverting the expression above delivers an expression for the ε^1 corresponding to the student who just marginally chooses not to drop j^1's section of course $\kappa = 1$.

$$\frac{s_{l^*} - \beta_{j^1} \cdot f_{j_1}(a_0, \mathbf{x}_1) - c_{j_1}}{\sigma^1_\varepsilon} = F^{-1}_\varepsilon\Big[1 - \mathbb{P}\big(R^1_{i,j^1} = 1, \mid a_0,\ \mathbf{x}_1,\ j^1\big)\Big]. \tag{2}$$

We can now make some progress by combining the identified expressions just derived. In particular we can use the latter to pin down the identity of the average marginal student relative to obtaining a score weakly above s_l. Once this is achieved, the same steps followed in the previous section can be used to infer the the variance parameters and the production function of at least one $\kappa = 1$ professor, namely the reference professor $\hat{j}_1$. Proposition 6 formally states and proves this claim.

Proposition 6. *The image $f_{\hat{j}_1}(a_0, \boldsymbol{x}_1)$ and the variance parameters $\sigma^1_\eta, \sigma^1_\varepsilon$ are point identified.*

Proof. Consider equation 1 describing the fraction of students (within the conditioning set) that obtains a score weakly above s_l in j^1's $\kappa = 1$ section. In particular, we consider the case for $l = l^*$, the cutoff above which students obtain a pass score. By algebraically manipulating this expression we obtain what follows,

$$\mathbb{P}\big(S^1_{i,j_1} \geq s_{l^*} \mid a^0,\ \mathbf{x}_1,\ j^1,\ R^1_{i,j^1} = 0\big)$$
$$= \int_{\varepsilon^1} \mathbf{1}\Bigg\{\varepsilon^1 \geq \underbrace{\frac{s_{l^*} - \beta_{j^1}\cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}}_{I}\Bigg\}\Bigg[1 - F_\eta\Bigg(\frac{\sigma^1_\varepsilon}{\sigma^1_\eta}\cdot\Bigg\{\underbrace{\frac{s_{l^*} - \beta_j\cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}}_{II} - \varepsilon^1\Bigg\}\Bigg)\Bigg] f_\varepsilon(\varepsilon^1) d\varepsilon^1.$$

Notice that terms I and II (both of which coincide) are quantities we have previously identified in equation 2. We can treat them as known quantities in the equation above. Since the left hand side is also an observed quantity (i.e., crucially, this is true because the researcher is capable of observing scores for students who don't drop the course), we can treat the identity above as just a function of the quotient $\sigma^1_\varepsilon/\sigma^1_\eta$. Furthermore, it is easy to see that under the region of integration considered, term II is always below ε^1 which implies that an increasing of the quotient $\sigma^1_\varepsilon/\sigma^1_\eta$ corresponds to a pointwise decrease of the integrand considered in the right hand side. It follows from standard inversion arguments that we can use the identity above to identify the true value of the quotient $\sigma^1_\varepsilon/\sigma^1_\eta$.

Let's now consider the analog to the previous expression for an arbitrary scores threshold s_l. This is given by the equation below,

$$\mathbb{P}\big(S^1_{i,j^1} \geq s_l \mid a_0,\ \mathbf{x}_1,\ j^1,\ R^1_{i,j^1} = 0\big),$$
$$= \int_{\varepsilon^1} \mathbf{1}\Bigg\{\varepsilon^1 \geq \frac{s_{l^*} - \beta_{j^1}\cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}\Bigg\}\Bigg[1 - F_\eta\Bigg(\frac{\sigma^1_\varepsilon}{\sigma^1_\eta}\cdot\Bigg\{\underbrace{\frac{s_l - \beta_{j^1}\cdot f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma^1_\varepsilon}}_{\theta(s_l \mid a_0, \mathbf{x}_1, j^1)} - \varepsilon^1\Bigg\}\Bigg)\Bigg] f_\varepsilon(\varepsilon^1) d\varepsilon^1.$$

It is clear from the preceding discussion that both terms $(s_{l^*} - \beta_j f_{j_1}(a^0, \mathbf{x}) - c_j)/\sigma^1_\varepsilon$ and $\sigma^1_\varepsilon/\sigma^1_\eta$ inside the integral term can be treated as known quantities. The key observation is then that we can treat the right hand side as a monotone function of the quotient $(s_l - \beta_j f_{j_1}(a^0, \mathbf{x}) - c_j)/\sigma^1_\varepsilon$ and $\sigma^1_\varepsilon/\sigma^1_\eta$ for any score cutoff s_l we entertain. We can then follow the same arguments as in section **??** by considering the system of equations defined by,

$$\theta(s_l|a_0, \mathbf{x}_1, j^1) = \frac{s_l - \beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma_\eta^1} \text{ and } \theta(s_{l'}|a_0, \mathbf{x}_1, j^1) = \frac{s_{l'} - \beta_{j^1} f_{j^1}(a_0, \mathbf{x}_1) - c_{j^1}}{\sigma_\eta^1}.$$

As before, it is easy to see that by considering $s_l \neq s_{l'}$, the system above has a unique solution in terms of quantities σ_ε^1 and $\beta_{j^1} \cdot f_{j^1}(a_0, \mathbf{x}_1) + c_{j^1}$. The former, together with our previous identification of the quotient $\sigma_\varepsilon^1/\sigma_\eta^1$, allows us to recover the variance term σ_η^1. In turn the result for $\kappa = 1$'s reference professor follows from our grading policy normalization of $(\beta_{\hat{j}^1}, c_{\hat{j}^1}) = (1, 0)$ ■

We can mimic the marginal logic student logic followed in Section **??** when trying to understand the content of Proposition 6. In doing so, it is useful to recall the main challenges arising from the possibility of student's dropping a course: (i) we don't observe the score of students who drop the course, (ii) the unobserved draw affects a student's incentive to drop the course. The first part of Proposition 6 shows we can easily correct for the first issue by using our observations of who students are dropping. In other words, we can infer who the marginal student dropping the course is and use this observation when constructing an identity describing the marginal student obtaining a score of s_l. The second part shows that even when ε draws affect the scoring equation, for our purposes we can focus in identifying the marginal student for a ε draw of zero.

The remainder of the argument tracks closely our previous work on Section **??** in that we used data for the student's performance on the second course of the sequence to disentangle $\kappa = 1$'s grading policies and production functions. For instance, consider all students in the conditioning set who after obtaining a pass score for $\kappa = 1$, enroll j^2's of $\kappa = 2$ under covariates $\mathbf{x}_2$. We are interested in an expression for the fraction of these students who obtain a score weakly above s_l. Our model implies the following expression,

$$\begin{aligned}
&\mathbb{P}\big(S_{i,j^1}^2 \geq s_l \mid a_0,\ \mathbf{x}_1,\ \mathbf{x}_2,\ j^1,\ j^2,\ R_{i,j^2}^2 = 0\ S_{i,j^1}^1 \geq s_{l^*}\big) \\
&\qquad = \int_{\varepsilon^2} \mathbf{1}\left\{\varepsilon^2 \geq \frac{s_{l^*} - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2}}{\sigma_\varepsilon^2}\right\} \\
&\qquad\qquad \times \left[1 - F_\eta\left(\frac{s_l - \beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \mathbf{x}_1), \mathbf{x}_2) - c_{j^2} - \sigma_\varepsilon \cdot \varepsilon^2}{\sigma_\eta^1}\right)\right] f_\varepsilon(\varepsilon^2) d\varepsilon^2.
\end{aligned}$$

Exactly the same arguments as in the preceding discussion can be applied to the $\kappa = 2$ problem. While direct inversion of the expression above is not possible, we can infer the primitives of interest by using our observations for how many students choose to drop

j^2's section. After achieving this the results are just as those considered in Section **??** in that we can identify $\kappa = 1$ production functions given injectivity of the $\kappa = 2$ instructor's production function. Below we state the result without a proof as it is identical to the arguments already outlined.

Proposition 7. *The following identification results hold,*

1. *The image of the composition* $\beta_{j^2} \cdot f_{j^2}(f_{j^1}(a_0, \boldsymbol{x}_1), \boldsymbol{x}_2) + c_{j^2}$ *and the variance term* $\sigma^2_\eta, \sigma^2_\varepsilon$ *are point identified,*

3. *Suppose that the learning production function* $f_{j^2}(\cdot\,, \boldsymbol{x}_2)$ *is injective. Then the image* $f_{j^1}(a_0, \boldsymbol{x}_1)$ *is point identified provided the existence of* $\tilde{a}_0$ *such that* $f_{j^2}(f_{j^1}(a_0, \boldsymbol{x}_1), \boldsymbol{x}_2) = f_{j^2}(f_{\hat{j}^1}(\tilde{a}_0, \boldsymbol{x}_1), \boldsymbol{x}_2)$.

As these parameters are not specific to individual professors, we provide detailed point estimates and standard errors for them in Table **??**. The interpretation of these parameters requires the consideration of the parameterization choice for the production function model. For instance, a negative value for $\mu_{0,l}$ or γ_l implies a decrease in the level of the learning production function, while a negative values for $\mu_{1,l}$ imply diminishing returns to increases in a student's ability.