What You Don't Know Might Deter You:
The Effect of Information Provision on Minority Retention in Undergraduate Economics

Anjali Pai

November 23, 2023

Note: I ran a second wave of this study in the spring of 2023, after the completion of this draft. As I am still collecting data on course enrollments and academic performance for the students in the second wave, those results have not been included in this version of the paper. An updated version with results from both waves of the study will be available at this link by January of 2024.

ABSTRACT

This study tests whether low-touch information provision can overcome barriers to entry in the undergraduate Economics major that may disparately impact underrepresented minority students. I run a large-scale field experiment in an introductory undergraduate microeconomics course that examines the effect of an email intervention containing information on the potential research topics, potential careers, expected income, and diversity of researchers in the Economics field. Through a combination of administrative data and survey-elicited beliefs, I am able to study short-run course performance and course selection, long-run majoring decisions, and effects on students' beliefs about the field. For underrepresented minority students, the information intervention increases the likelihood of enrollment in the next Economics course by 12.3 percentage points. Examining this result further, I find strong suggestive evidence that the intervention motivates lower-performing students to enroll in a subsequent Economics course. Finally, I examine the belief updating mechanisms underlying these results. Survey data shows that underrepresented minority students increase their awareness of unconventional research topics associated with Economics with no significant shifts in the other domains of information relevant to the treatment. These results are consistent with the theory that information interventions may be most effective for the "marginal" student who is interested in, and capable of succeeding in, Economics, but in the absence of full information may be deterred by initial low performance or rigorous courses. [1]

[1]Thank you to the UCSD Department of Economics' *Diversity in Economics Research Fellowship* for providing the funding for this project. A major thank you to Melissa Famulari for making this study possible and for providing data, implementation support, and guidance throughout. Thank you to Cindy Lyons, Pegah Parsi, and David Garrison from the Office of the Registrar at UC San Diego, for their partnership and ongoing access to administrative data. I have also benefited from comments and seminar feedback from professors and graduate students at the UC San Diego, and attendees at the 2022 All California Labor Economics Conference. Finally, I owe a debt of gratitude to my advisor, Julian Betts, for his support throughout this project.

1 Introduction

STEM fields have experienced a troubling pattern in recent years: despite high job security, high pay, and high long-run returns, diversity in these fields has remained low. The characteristics of STEM jobs should make them highly attractive to students from all backgrounds, yet students of various non-White racial backgrounds have historically been underrepresented in high-paying fields such as computer science, engineering, and economics. While the share of bachelor's degrees awarded to Black or African American students in the U.S. has been steadily increasing in the past 25 years (from 7.2% in 1995 to around 10.2% in 2020), the share of STEM and economics degrees have stagnated. Over this same period, the fraction of Black students receiving STEM undergraduate degrees increased slightly from 6% to 6.5%, while the fraction receiving economics undergraduate degrees *decreased* slightly from 6.4% to 5.2% (Hoover and Washington, 2021). While the fractions of Hispanic individuals receiving STEM and Economics degrees have increased over this period (from around 5% in 1995 to around 13.7% for both categories), these levels are still well below the fraction of Hispanic individuals in the population.

The gap between the fraction of STEM-related degrees awarded to minorities and the fraction of non-STEM bachelor's degrees captured by these groups is not inherently inefficient. If this gap is driven entirely by differences in major and career preferences across groups, then the current allocation of students across degrees may maximize each individual student's utility. On the other hand, if the gap is caused by supply-side or demand-side frictions that adversely impact minority populations, the current allocation may be inefficient both for individuals and society more broadly. This paper focuses on one potential demand-side friction that could contribute to the representation gap: misinformed or inaccurate beliefs about STEM fields amongst students from minority racial backgrounds. Specifically, if students from minority backgrounds lack pre-college exposure to STEM subjects, they may not be fully or correctly informed about the returns to these fields. Furthermore, stereotypes about the types of jobs associated with these fields or the types of individuals who succeed in them could dissuade uninformed students from pursuing these subjects.

If incomplete information is a barrier to persistence in STEM-related fields, one potential solution is an information intervention targeted to address the misconceptions students may have. I examine this theory through a large randomized controlled trial in an undergraduate economics course at a public university. The experiment tests whether an email providing information about research topics, careers, potential future income, and diversity in the field of economics affects short- and long-term course enrollment, performance, and majoring decision for underrepresented minority (URM) students. I also administer baseline and endline surveys to identify which (if any) dimensions of the information intervention shift students' beliefs. In all analyses, I follow the definition of "underrepresented minority" used by the American Economic Association, which includes individuals who identify as Black, Hispanic, Native American/Indian American, or Native Hawaiian/Pacific Islander.

This low-touch intervention increases the likelihood of enrollment in a subsequent economics course for URM students by 12.3 percentage points. Moreover, the treatment shifts the composition of students enrolling in a subsequent economics course by inducing enrollments from lower-performing URM students. This result is consistent with the theory that students who lack prior exposure to economics, (and therefore are more likely to be influenced by the intervention), may also lack the necessary preparation for success in early courses of the economics major. Examining belief updating, I find that URM students primarily shift their beliefs about the scope of social issues encompassed in economics research, rather than about expected income or career-options. This suggests that the mechanism driving the increase in enrollment for lower-performing URM students may be an increased "interest" in economics due to new information about the breadth of the field. This suggests that information provision has the potential to counteract a cycle where minorities who enter college with lower academic preparation and lower exposure to STEM-related fields fail to remain in these fields long enough to build the skills and tools necessary to succeed.

Economics, while not explicitly a STEM field, is a comparable context for studying incomplete information. Economics has seen greater underrepresentation of students from certain racial and ethnic groups than STEM subjects in the past two decades. For example, only 5.2% of economics bachelor's degrees were awarded to Black students in 2020 as compared to 6.5% of STEM degrees (Hoover and Washington, 2021). Moreover, there are reasons to expect that URM students may have lower pre-college exposure to economics than non-URM students. In order to understand why this could be the case, consider two potential channels of pre-college exposure to fields of study for students. The first channel is through family and community networks. Given historic underrepresentation of minorities in economics degrees and careers, minority communities may be less likely than non-minority communities to include individuals with degrees or careers in economics. In this case, community networks may not provide URM students with information about the scope of, and returns to, an economics major. A second channel is through school curricula and courses. While there is limited evidence about the economics course-offerings in high school, 2019 data from the College Board suggests that only 24% of schools with Advanced Placement courses offer AP Macroeconomics, and 20% offer AP Microeconomics. In comparison, Advanced Placement courses in subjects such as biology and chemistry are offered at close to 50% of schools, while courses in English and History are offered at nearly 60% of schools (CollegeBoard, 2019). The lower prevalence of AP courses in economics suggests that they may not be a top priority for schools with Advanced Placement offerings. Although this does not necessarily imply that URM students would have lower exposure to economics in schools than non-URM students, it does suggest that economics may not be viewed as a "core" subject that students should be exposed to prior to college. If low-income schools have staffing and funding constraints that limit the number of advanced courses they can offer, they may prefer to offer the "common" AP courses rather than the less common ones.

This paper contributes to a growing literature on the supply- and demand-side determinants of major choice and persistence in STEM majors. On the supply side, there is evidence that demographics of instructors and role models may improve early-pipeline retention of minorities and women (Carrell et al. (2010), Hale and Regev (2014), Fairlie et al. (2014), Kofoed et al. (2019)). Related work shows that implicit and explicit biases affect how instructors treat students from minority backgrounds (Carlana, 2019). On the demand side, early papers suggested that ability-sorting in higher education was primarily driven by heterogeneous preferences across majors (Arcidiacono (2004), Beffy et al. (2012)). More recent work from Wiswall and Zafar (2015) provides evidence that students' beliefs may play a larger role in major choices than previous work had estimated. Wiswall and Zafar conduct a large-scale information experiment where they test whether students update their beliefs about own-ability, future income, and long-run returns for a large set of college majors in response to information about those fields. They find that students tend to misestimate earnings at the baseline and update their beliefs about the average earnings for each field in response to new information. In line with the prior literature, however, they also find that while beliefs about earnings and own-ability are important for major choice, these factors are dominated by heterogeneous preferences and tastes. My work adds to this literature by studying how field-specific information that extends beyond expected earnings influences students' belief-updating and major selection.

This paper also contributes to the literature on interventions to improve enrollment and retention in STEM fields. The closest study in this literature is Bayer et al. (2019), which tests whether providing information about the breadth and diversity of economics to incoming college freshmen at a sample of liberal arts colleges affects short and long-run course enrollments. They find that the intervention increases the probability of taking and completing a first semester Economics course by 3 percentage points across all students and by nearly 11.4 percentage points for first generation college students; however they find no evidence that these effects persist after the initial course. These results suggest that information provision may be helpful in inducing *initial enrollment* in economics courses, but highlights that retention is still a concern. The current study complements the work by Bayer et al. by providing evidence on the effect of a low-touch intervention with similar types of information on retention of students who have expressed preliminary interest in the study of economics. The combination of my results and the positive effects identified by Bayer et al. suggest that ongoing information provision at various stages of undergraduate education may be more valuable than a one-off intervention. Related work from Chambers et al. (2021) finds suggestive evidence that an email or series of texts containing links to videos and infographics about economics leads to small but positive increases in students' self-reported probability of taking a subsequent economics course; however, they find no effect on actual course enrollment. Examining a more intensive intervention, Porter and Serra (2020) finds that providing information through female alumni mentors increases female students' probability of taking a subsequent Economics course by nearly 100%; however, they are unable to disentangle the effect of having female mentors from the effect of the information provided by those mentors. While the population I study is more comparable to the latter two papers, I would expect my effect sizes to be closer to those of the Bayer et al., since their intervention design and delivery model are most similar to mine.

The remainder of this paper is structured as follows. Section 2 proposes a theoretical model for student major choice and provides motivation for why information interventions may have greater impact for URM students; Section 3 discusses the experimental design; Section 4 presents descriptive statistics for the experiment sample; Section 5 shows results for course performance, course enrollment, and majoring behavior; Section 6 explores effects on belief updating; and finally, Section 7 discusses the policy implications of these results and concludes.

2 Theory

In order to provide some intuition about the conditions under which information provision may affect student major decisions, this section examines a generalized model of major choice. Appendix C contains a more detailed outline of this theory.

Consider an undergraduate student who is selecting a major, m from a set of N majors: $m \in \{m_1, m_2, \ldots, m_N\}$, where the utility from major m is a function of long-run income (I_m) and long-run satisfaction (S_m).[2] We can define a student's expected utility from major

[2]Note that satisfaction is a broad term that can include satisfaction attained from jobs that result from

m as

$$U_m = F(\mathbb{E}[I_m], \mathbb{E}[S_m])$$

where $\mathbb{E}[I_m]$ and $\mathbb{E}[S_m]$ are the student's expectations of the level of income and satisfaction they will achieve by majoring in m. Define the true utility of major m as

$$U^*_m = F(I^*_m, S^*_m)$$

where I^*_m and S^*_m are the *true* levels of income and satisfaction the student would attain from pursuing major m. Note that if a student has perfect information about major m, $\mathbb{E}[I_m] = I^*_m$, $\mathbb{E}[S_m] = S^*_m$, and $U_m = U^*_m$.

Without loss of generality, assume the student has a strict preference ranking over majors defined by:

$$U_{m_N} > U_{m_{N-1}} > \cdots > U_{m_2} > U_{m_1}$$

Now, consider information provision on some major j. For simplicity, assume that information provision moves the student to perfect information about major j. This means that after receiving information the student will update their expected utility for major j from U_{m_j} to $U'_{m_j} = U^*_{m_j}$. Holding the expected utilities of all other majors constant leads to three key results:

1. If $U_{m_j} = U'_{m_j} = U^*_{m_j}$ information provision will **not affect** the student's preference ranking over majors.

2. If $U^*_{m_j} < U_{m_N}$ information provision will **not affect** the student's ultimate major choice. In this case, the student is no worse off from receiving the information than they would have been otherwise

3. If $U^*_{m_j} > U_{m_N}$ information provision will cause the student to switch to preferring major j over major N (their baseline preference). This switch will be utility-increasing if $U^*_{m_j} > U^*_{m_N}$ and utility-decreasing if $U^*_{m_j} < U^*_{m_N}$

Result 1 gives the fairly obvious condition that in order for information provision on major j to shift a student's major preferences, they must lack information about major j at baseline. Result 2 points out that in the case where a student is misestimating their expected utility across *all majors*, utility provision on a single major will not necessarily

majoring in m as well as the satisfaction gained from the major itself (i.e. coursework, community, passion for the subject).

shift the student's preferences across majors. Finally, Result 3 crucially notes that in the case where information provision on a single major actually shifts preferences over majors, it is ex-ante ambiguous as to whether that major shift will be utility increasing or decreasing.

For the purposes of this paper, Result 1 is the most important. Since the outcomes studied in this paper are limited to educational outcomes within 1-2 years of the intervention it is impossible to identify the long-run utility-impacts of this intervention. I leave that to future work in this area. Instead, this paper seeks only to answer the question of whether a low-touch information intervention can affect students' behavior and major decisions. Result 1 provides the intuition that if this information intervention affects students' calculation of the expected utility from majoring in Economics, that effect should be stronger for those students that lack the most information at baseline.

Building on this observation, I hypothesize that information provision will be most effective for URM students and students from low socioeconomic backgrounds for the exact reason that these students are more likely to have a larger lack of information at baseline. One potential source of this lack of information is that individuals from minority racial backgrounds may have less exposure to Economics prior to entering college. It is reasonable to believe that an individual's exposure to certain career paths will be directly affected by the extent to which that career is represented within their immediate community. If individuals' communities are shaped by their racial identity, then we might expect that individuals from racial communities that are historically underrepresented in Economics may have less exposure to Economics. This may cause them to have less information about the field at baseline relative to their non-URM counterparts. If this is the case, we may expect that information provision on Economics may be more effective for URM and lower socioeconomic individuals who are more likely to lack prior exposure to the field. This intuition also implies that in general, information interventions may have the greatest impact in fields that posses larger diversity gaps.

3 Experimental Design

To test these hypotheses, an experiment was designed and conducted in a Principles of Microeconomics class at a large research university in the Fall academic term of 2021.[3] To supplement this original sample and address the external validity of first round results, a second wave of the experiment was run in the Spring academic term of 2023, in the same

[3]While this study was not pre-registered, the design and subgroup analyses conducted were decided prior to applying for IRB-approval and were not modified after receiving approval.

course with a different instructor and smaller sample of students. Heretofore, I will refer to the fall 2021 experiment as “Wave I” and the spring 2023 experiment as “Wave II”.

The experiment tested the effects of an email containing information about the scope of topics that can be studied in Economics, potential careers attainable with an undergraduate degree in Economics, and expected earnings from an undergraduate degree in Economics. The information treatment also highlighted three diverse researchers in Economics. The sections on topics in Economics and diverse researchers follow the structure used by Bayer et al. (2019); however the specific format of my information intervention, as well as the sections on income and careers, are novel to this study.

All students enrolled in the course were informed of the study at the beginning of the academic term and given access to a form to opt out of the study at any point during the academic term. Students were informed that if they agreed to participate in the study, they would be given the opportunity to complete two short surveys for a small number of extra credit points, (equivalent to around 2% of their final course grade).[4] Students were *not* told that they would be assigned to treatment and control groups or that they would be sent information as part of the study. The information intervention was sent approximately halfway through the academic term in the form of an email from a professor serving as the Vice-Chair of Undergraduate Economics. One goal of sending this email through the Vice-Chair was to increase the credibility of the information provided and boost the likelihood that students would read the email.

A second goal of sending the email through an individual *unaffiliated* with the course in which the experiment was run was to mitigate experimenter demand effects. This is a common concern in experiments wherein participants in an experiment adjust their behavior based on inferences they make about the goals of the study or the experimenter's hypotheses (De Quidt et al., 2019). Since the behavior studied in this experiment involves large-stake decisions like course enrollment and majoring, experimenter demand effects are less of a concern for these outcomes; however it may be a concern for students' endline survey responses. If students know that this study involves an information intervention designed to shift their beliefs about Economics, they may pay greater attention to the information intervention, or adjust their answers to the endline survey in a manner that does not reflect how they would act in a non-experimental setting. Sending the information intervention through an administrator decreases the likelihood that students would associate the treatment with the study being run in their Principles of Microeconomics course. Similarly, not directly informing stu-

[4]In Wave I, students were offered 2 extra credit points for opening each survey and 8 additional extra credit points for completing the survey. The final course grades were determined out of 1000 points, of which the surveys made up 20 total points.

dents that they will be receiving an information intervention places greater distance between the email they receive and the surveys they are responding to as part of this "study".

3.1 Baseline Survey and Randomization

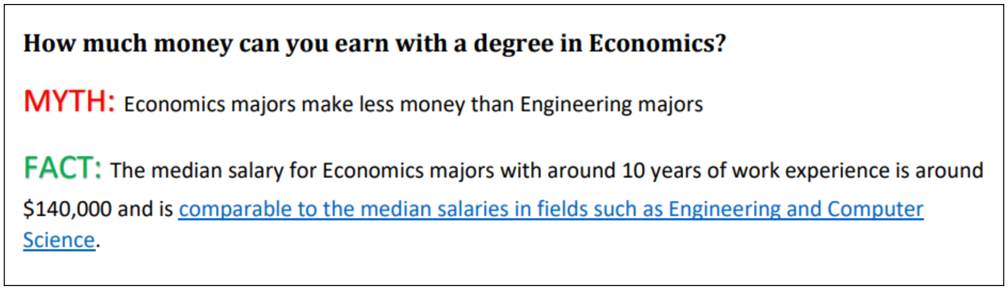
All students enrolled in the course were sent a baseline beliefs survey at the start of the second week of the academic term. When the baseline survey was released, 980 students were enrolled in the course. At the conclusion of the survey period (nine days after the survey was released) there were 940 students enrolled in the course, of whom 817 completed the baseline survey.

The baseline survey elicited students' beliefs on a few key dimensions: the types of research associated with Economics, the types of careers an Economics major would be qualified for, and the level of expected income for an Economics major compared to other close major-substitutes. The survey also asked for a set of demographic characteristics including: gender, ethnicity, highest degree attained by a parent, international versus domestic student status, exposure to Economics prior to college, and affiliation of a parent with the field of Economics.

The responses from this baseline survey were used for a stratified randomization design that randomized students within multi-dimensional blocks defined on four demographic characteristics: gender, URM-status, international status, and highest parental education. Stratifying increases the chances of balance between the treatment and control for relatively small groups that may also be the most likely to benefit from the intervention. URM students are a particularly good example. Due to the aforementioned "diversity gap" that motivates this study, we may expect URM students to make up a relatively low fraction of the sample; however we also expect the information treatment to have larger effects for these students. Stratifying on this characteristic precludes finite-sample bias which, in an extreme case, could lead to no URM students in the treatment group despite an overall large sample. The other strata are similarly motivated.[5]

Practically, gender, URM-status and international status were defined as dummies where 1's represent being female, URM, and international respectively. Highest parental education was defined as three bins - Less than a four-year degree, four year degree, and post-graduate degree. All four strata contained a separate category for "missing" responses leading to three categories for non-cisgender male, underrepresented minority, and international student, and

[5] Including international-status addresses an issue specific to the Economics department at this university, where nearly 40% of students in Intermediate Microeconomics are international.


How much money can you earn with a degree in Economics?

MYTH: Economics majors make less money than Engineering majors

FACT: The median salary for Economics majors with around 10 years of work experience is around $140,000 and is comparable to the median salaries in fields such as Engineering and Computer Science.

Figure 1: Example of one of the sections in the treatment email. Blue underlined text represents a hyperlink to a relevant American Economic Association page

four categories for parental education. The final randomization occurred within 108 cells defined on the intersections of these four variables.

3.2 Treatment and Active Control

This experiment tests the effects of an email containing information on four primary dimensions: (1) the scope of topics that can be studied in Economics; (2) potential careers attainable with an undergraduate degree in Economics; (3) expected earnings from an undergraduate degree in Economics; and (4) short profiles of three diverse researchers in Economics. The information treatment in this study was designed to directly address common misconceptions about Economics and highlight diversity in Economics research. Previous studies such as Bayer et al. (2019) have focused on emphasizing diversity in the field. The treatment in this study distinguishes itself from previous work in three key ways. First, the design is structured around identifying common "myths" about Economics and contrasting them with "facts". Figure 1 shows an example of this structure. Second, unlike previous papers, this intervention expands beyond just offering information about the range of research topics and the highlighting of diverse researchers, and includes information on the range of careers and potential expected income as well. Finally, the intervention more explicitly lists examples of research topics and careers, before directing students to the AEA website for further information. Appendix B shows the exact contents of the treatment and control emails.

The control group in this study receives an email of similar length and structure to the treatment email. The primary goal of this is to mitigate experimenter demand effects, a phenomenon in experiments where participants may shift their behavior in response to inferences about the type of behavior the experimenter is attempting to elicit or study. This is of concern if inferences about the goals of the study lead treated students to shift their behavior in the same direction as the predicted treatment effects. In this experiment, the primary concern is that experimented demand effects may bias the estimation of belief updating; specifically, if students infer that the goal of the treatment is to shift their beliefs about Economics, they may respond differently in the endline beliefs survey than they did on the baseline.

One method of addressing this is to use an "active control", which is a control treatment that at the looks similar to the information treatment but contains only benign information that should not affect behavior (Haaland et al., 2020). The active control for this study uses the same format as the treatment email: in particular, it includes the same information about Economics department resources and university-wide resources at the start, and mimics the use of pictures alongside text in the body of the email.[6] To avoid providing the control group with information that could sway their decisions, the control email contains information about the content of the Principles of Microeconomics course drawn directly from the course syllabus, which was already available to students at the time of the study.

3.3 Endline Survey

Approximately two weeks after students received the email containing the information treatments, all students in the course were sent an endline survey. Of the 940 students enrolled in the study sample, 822 responded to the endline survey. The first part of the endline survey contained the same belief-elicitation questions that were included in the baseline survey. Repeating the same set of questions allows for clean identification of belief-updating due to the information intervention. The second part of the survey contained a set of questions gauging students' engagement with the actual information treatment. Since I am unable to identify if students *actually read* the emails they were sent, these endline survey questions are the next best option for estimating a lower bound on the number of students who engaged with the treatment. Students were asked whether they recalled receiving an email from the Vice-Chair of Undergraduate Economics and what specific contents they recalled from that email. The former question attempts to identify a lower bound on "viewing" of the email while the latter attempts to identify "engagement" with various part of the information treatment.[7]

The last part of the endline survey contains a series of questions gauging students' "social

[6] Since the use of images in emails is not too common, this is one aspect of the treatment email that may draw students' attention and lead them to infer that they are in the "treatment group". The use of pictures in the control email as well helps mitigate that concern.

[7] Since students are not incentivized to report truthfully and are asked to draw from their "recollection", the measure of interaction derived from the endline survey will likely be noisy. For this reason, the primary results in this paper are intent-to-treat estimates; in future robustness checks I plan to use this data to estimate treatment-on-treated effects.

connections" within the course. These questions are designed to help identify treatment group spillovers. Although the active control group should mitigate spillovers, spillovers are impossible to rule out given the large difference in the kinds of information provided in the treatment and control emails. In order to predict the likelihood of such spillovers, students were asked about the frequency at which they attend in-person classes, in-person TA sessions, and in-person review sections, which are all locations in which they would be more likely to interact with other students in their course. In addition, students were directly asked how often they work with other students in their course, how many other students they interact with, and how often they discuss topics such as major choices and diversity and representation with their friends.

3.4 Administrative Data

The two surveys run in the context of this course allow me to identify effects of the treatment on belief-updating. In order to attain data on students' course grades, subsequent course enrollment, and majoring behavior, I partnered with the Registrar's Office. Through the Registrar's Office I have access to the following information for all students in this study sample, for the entire academic year following the intervention: the list of Economics courses taken, final course grades for each Economics course, and current declared major at the start and end of each academic term. For Principles of Microeconomics, in addition to the final course letter grades from the registrar, I have access to students' raw final point totals, and the scores from two midterm exams. The timing of the treatments was such that the treatment emails were sent out *between* the administration of the two midterms. Therefore the scores from the first midterm can be used as a metric for baseline performance.

4 Data

4.1 Balance

I begin by verifying the validity of the randomization. Table 1 shows the balance on baseline characteristics between the final treatment and control groups. We can see from this table that the four characteristics that were blocked on for randomization are balanced by design. What is worth noting is that almost all baseline characteristics that were not actively balanced in the randomization are still balanced between the treatment and control groups. The p-values in Column 3 reflect that for almost all baseline covariates, a two-sample t-test rejects any difference between the treatment and control means. This gives reasonable confidence that the treatment is in fact randomly assigned.

Table 1: Balance on Baseline Demographics by Treatment Group

	Control	Treatment	t-test(p)
Strata			
Female	0.403	0.400	0.92
URM	0.181	0.179	0.93
International	0.312	0.303	0.78
Bins of Parental Educ.	1.160	1.161	0.98
Other Covariates			
College Start Year	2020.72	2020.70	0.71
Parent is Professor/Researcher in Econ	0.019	0.013	0.43
Economics Exposure			
Exposed to Econ Before College	0.585	0.539	0.16
Exposed to Econ at College	0.062	0.051	0.48
Entered Econ Major	0.353	0.313	0.20
Baseline Course Data			
Total Survey Bonus Points	17.242	17.541	0.42
Section A	0.303	0.392	0.00
Section B	0.348	0.354	0.86
Section C	0.346	0.247	0.00

The one exception worth discussing here is the course section. The Principles of Microeconomics class in which this experiment was run consisted of three sections. Since this experiment was run during the second year of the COVID-19 pandemic, the university was operating under a partially in-person, partially remote course structure. Of the three sections, sections 1 and 2 were in-person and section 3 was fully remote (and practically more difficult). Table 1 reveals an idiosyncratic imbalance in the course section distribution across the treatment and control groups. One potential explanation for this imbalance is that the attempt to *force balance* on a set of demographic characteristics inadvertently led to an imbalance in the course section distribution due to the demographic ratios present in each course section. To assess this, Table 2 shows the means and standard deviations of each baseline variable for each of the three course sections. One initial observation from this table is that there seems to be a slight gender imbalance across the three sections. Compared to Section B, which has a similar fraction of women to the pooled sample (40%), Section A has a fewer women (38%) and Section C has more women (42%). One theory for the section imbalance between the Treatment and Control groups is that forcing balance on gender across all sections in the presence of these gender ratio differences may have inadvertently forced an imbalance of treatment and control allocations within each section.

Table 2: Means of demographic covariates by course section

	(1) Section A	(2) Section B	(3) Section C	(4) Pooled
Treatment	0.558	0.513	0.420	0.501
	(0.498)	(0.501)	(0.495)	(0.500)
Female	0.379	0.401	0.424	0.402
	(0.486)	(0.491)	(0.495)	(0.491)
URM	0.182	0.199	0.155	0.181
	(0.387)	(0.400)	(0.363)	(0.386)
International	0.290	0.339	0.282	0.303
	(0.455)	(0.474)	(0.451)	(0.460)
Bins of Parental Educ.	1.104	1.231	1.151	1.159
	(0.799)	(0.806)	(0.797)	(0.803)

Standard errors in parentheses

To explore this possibility further, we can look at the balance of baseline covariates between the treatment and control groups *within* each section. I focus specifically on Section A (Table 3) and Section C (Table 4). In line with the hypothesis that the treatment group imbalance may be driven by course section gender imbalances, Table 3 reveals that within Section A, the treatment group is 44% female compared to the treatment group which is 30% female. Looking instead at Table 4, the female ratio in the treatment and control groups is almost the exact reverse of the ratios in Section A. Unsurprisingly, Appendix Table A.1 reveals balance in the gender ration across the treatment and control groups within Section B.

Table 3: Balance of covariates for course Section A

	Control	Treatment	t-test(p)
Strata			
Female	0.300	0.442	0.02
URM	0.183	0.184	0.99
International	0.261	0.305	0.42
Bins of Parental Educ.	1.167	1.071	0.33

Absent an obvious reason for these section imbalances, there is minimal concern that the randomization itself is invalid. Preliminary results from a second wave of this study reveal balance on all baseline variables, suggesting that the imbalances in this sample are likely driven by pure chance. Regardless of their source, I account for any baseline imbalances by including course section dummies, baseline demographic controls, and baseline performance controls in all my preferred specifications.

Table 4: Balance of covariates for course Section C

	Control	Treatment	t-test(p)
Strata			
Female	0.489	0.333	0.02
URM	0.137	0.176	0.40
International	0.282	0.284	0.96
Bins of Parental Educ.	1.156	1.168	0.91

4.2 Attrition

The other threat to internal validity would be differential attrition. At the time of randomization and treatment assignment, 940 students were enrolled in the course. As of the last day of the endline survey period, only four students had dropped out of the course leading to very low concerns about differential course attrition. A bigger concern would be differential attrition in study participation — specifically, attrition between the baseline and endline surveys. 817 students responded to the baseline survey and 822 responded to the endline survey. Of these 822 endline survey responses, only 752 were able to be matched to a baseline survey. Since belief updating can only be identified for students who have both a baseline and endline survey response the act of filling out a baseline survey but not an endline survey can be viewed as a form of attrition. Students were not told that these surveys were specifically "baseline" and "endline" surveys, nor were they provided any added incentive for taking both surveys; therefore there is no reason to believe that differential survey attrition would be an issue. Table 5 shows no differential attrition out of the course or between surveys.

Table 5: Attrition rates by Treatment Group

	Control	Treatment	t-test(p)
Dropped Class	0.002	0.006	0.32
No Endline Survey	0.079	0.060	0.24

It is also valuable to verify that attrition is not being driven by students of certain demographic characteristics. Table 6 shows the balance of baseline demographic characteristics between the sample of students with both baseline and endline responses, compared to the sample of students with only baseline responses. Based on the two-sample t-tests, there is no evidence of differential attrition by baseline observables.[8]

[8] The one exception is that there seems to be *less attrition* between surveys for students in Section C, although this may partially be due to the lower take-up of the baseline survey in this section compared to other sections. Again, this imbalance is unlikely to affect enrollment outcomes but could potentially affect the beliefs analysis.

Table 6: Balance of demographic characteristics by Attritted and Non-Attritted Samples

	Both Surveys	Baseline Only	t-test(p)
Strata			
Female	0.409	0.313	0.13
URM	0.180	0.175	0.91
International	0.307	0.313	0.93
Bins of Parental Educ.	1.170	1.048	0.24
Other Demographics			
College Start Year	2020.72	2020.64	0.28
Economic Exposure			
Parent is Professor/Researcher in Econ	0.017	0.000	0.29
Exposed to Econ Before College	0.557	0.631	0.25
Exposed to Econ at College	0.056	0.062	0.86
Entered Econ Major	0.335	0.308	0.65
Course Data			
Section A	0.343	0.415	0.24
Section B	0.352	0.338	0.82
Section C	0.305	0.185	0.04

4.3 Summary Statistics

Table 7 shows basic summary statistics for the experiment sample. Approximately 40% of students enrolled in Principles of Microeconomics identify as female. As predicted, the share of URM students is fairly low in the sample (only around 18%). Similarly, only around 22% of the sample consists of first generation students (defined as students who report that the highest education level of one of their parents is below a college degree). 30% of students are international and most students are in their first or second years of college. A final observation of note from this table is that around 33% of the students in this course entered college with a declared Economics major, and around 56% had some form of exposure to an Economics course prior to college.

Table 7: Summary Statistics for Baseline Demographics

	Mean	St. Dev.	Min	Max
Strata				
Female	0.402	0.491	0	1
URM	0.180	0.384	0	1
International	0.307	0.462	0	1
Bins of Parental Education	1.161	0.802	0	2
Other Demographics				
First Generation	0.218	0.413	0	1
In Fraternity/Sorority	0.036	0.187	0	1
In a Team Sport	0.092	0.289	0	1
College Start Year	2020.72	0.610	2016	2021
Economics Exposure				
Parent is Professor/Researcher in Econ	0.016	0.126	0	1
Exposed to Econ Before College	0.562	0.496	0	1
Exposed to Econ at College	0.057	0.231	0	1
Entered Econ Major	0.333	0.472	0	1

In addition to overall summary statistics, we may be interested in examining some summary statistics specific to URM students. Table 8 shows the means across the same set of demographic characteristics as before, but compared across URM and non-URM students. Here there are some meaningful differences that provide some added motivation for why URM students may be most likely to have a lack of information about Economics at baseline. Specifically, Table 8 reveals that URM students are, on average, more likely to be first generation, and slightly more likely to be female. Looking at the "Economic Exposure" measures, URM students are less likely to have a parent working as a researcher or professor in Economics and less likely to have taken an Economics course prior to college. These differences provide meaningful evidence that URM students may be less likely to be exposed to Economics prior to college. This lack of exposure may be due to lower socioeconomic conditions, which may be correlated with limited Economics course offerings in high school; however it may also be due to a lack of exposure to individuals working in Economics which is a possible result of historic underrepresentation in the field.

Table 8: Summary Statistics for URM vs Non-URM Students

	Non-URM	URM	t-test(p)
Female	0.388	0.458	0.12
International	0.358	0.069	0.00
Bins of Parental Educ.	1.258	0.685	0.00
First Generation	0.189	0.556	0.00
UCSD Start Year	2020.7	2020.6	0.01
Parent is Professor/Researcher in Econ	0.021	0.000	0.08
Exposed to Econ Before UCSD	0.667	0.569	0.03
Exposed to Econ at UCSD	0.070	0.042	0.21
Entered Econ Major	0.362	0.319	0.33

5 Results: Behavioral Outcomes

The design of this study allows for simple identification of the effects of information on a set of student performance and enrollment outcomes. The analysis for each outcome begins by looking at the overall treatment effect in the pooled sample before exploring heterogeneity by URM-status and socioeconomic-background (captured by an indicator for first-generation status). The hypotheses tested in this section are as follows:

H1: The treatment will improve immediate performance in the academic term of the intervention by increasing some students' interest in pursuing Economics, and thereby inducing an increase in immediate effort in their current Economics course.

H2: The treatment will increase the likelihood of course enrollment in the term following the intervention for those students whose interest in pursuing Economics was improved by the intervention.

H3: If treatment effects persist long enough to increase students' likelihood of majoring in Economics, these effects may show up in the year following the intervention.

Thinking specifically about the populations this intervention is targeted at, the theoretical underpinnings of this study predict that the effects of information provision should be largest for sub-groups of the population that have lower exposure to Economics. This lends itself to the following hypotheses regarding treatment-effect heterogeneity:

H4: The treatment effect will be stronger for URM students than non-URM students due to lower exposure to Economics prior to entering college.

H5: The treatment effects will be stronger for first generation students for similar reasons of lower exposure.

H6: If there is heterogeneity by gender, those effects will go in the direction of stronger treatment effects for women, since women may have stronger preferences for non-traditional careers and research areas that the intervention emphasizes.

5.1 Effect on Fall Course Performance

The first outcome of interest is course performance in the Principles of Microeconomics course where the intervention was run. If the information provided was effective in piquing

URM students' interest in the Economics major, we may expect to see an *improvement in course performance* driven by increased motivation to succeed in Economics and subsequent increased effort. Since the intervention was run approximately halfway through the academic term, there would have been sufficient time for students to adjust their effort and impact their final course performance.

Table 9 shows the treatment effects for URM students on midterm and final course grades in Principles of Microeconomics. The treatment effect for interacted students should be interpreted as the sum of the coefficients on the *Treatment* and *Treatment X URM* variables. The F-statistics and p-values reported at the bottom of the table come from a test that the sum of those two coefficients is different from zero. The table reveals that after controlling for the full set of baseline demographics[9] and baseline course performance[10] there is no significant treatment effect on either the second midterm or final course grades for URM students in this sample.

[9] For this, and all future models, "Demographic Controls" include: URM-status, gender, international-status, dummies for highest parental education (Bachelor's and Postgraduate), a dummy for entering the university as an Economics major, dummy for having a parent in an Economics-related career, a dummy for exposure to Economics prior to college, and dummies for the course section in the term of the intervention.

[10] I measure baseline course performance using students' raw grades on the first midterm, which occurred prior to the intervention. It is worth noting that due in large part to the section imbalances between the treatment and control groups, despite being a pre-intervention metric, there is a "treatment effect" on midterm 1 grades. Since this measure varies significantly between the treatment and control groups, I control for baseline performance in all first quarter grades regressions and use Midterm 2 grades (which are independent of Midterm 1 grades) as a post-intervention grades outcome alongside final course grades.

Table 9: Treatment Effect on Principles of Microeconomics Performance for URM Students

	Midterm 2			Final Course Points		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment x URM	-0.640***	0.188	0.0514	-121.2***	33.28	8.116
	(0.127)	(0.180)	(0.138)	(18.84)	(26.92)	(15.93)
Treatment	0.162**	0.0240	-0.0580	44.73***	14.05	-1.040
	(0.0690)	(0.0699)	(0.0577)	(10.68)	(10.20)	(7.237)
URM		-0.607***	-0.203*		-113.6***	-39.15***
		(0.140)	(0.108)		(21.39)	(14.33)
Constant	0.0389	-0.122	0.0215	755.4***	717.0***	743.4***
	(0.0481)	(0.107)	(0.0833)	(8.149)	(16.66)	(10.78)
Demographic Controls	No	Yes	Yes	No	Yes	Yes
Baseline Performance	No	No	Yes	No	No	Yes
F-Stat for Interaction	14.32	1.600	0.00260	15.66	3.550	0.227
p-Value for Interaction	0.000166	0.206	0.959	0.0000825	0.0599	0.634
Observations	797	785	785	797	785	785
R-squared	0.0339	0.135	0.447	0.0535	0.253	0.673

Standard errors in parentheses. Observations are at the individual level

F-statistics and p-values come from a t-test that *(Treatment x URM) + (Treatment) = 0*

* p<0.1, ** p<0.05, *** p<0.01

5.2 Effect on Course Enrollment One Quarter After the Intervention

A primary question in this study is whether such a low-touch intervention will have any effect on students' course enrollment and majoring decisions. To assess this, I look first at course enrollment in the academic term following the intervention. Of the students who enrolled in an Eocnomics course on quarter after the intervention, close to 70% enrolled in Principles of Macroeconomics. This is sensible given that Principles of Microeconomics and Principles of Macroeconomics are the two "lower-division core courses" necessary for the Economics major. In addition, in many cases these courses are prerequisites for "intermediate" courses in the major. Therefore it is reasonable to begin this analysis by looking at the decision to enroll in Principles of Macroeconomics one quarter after the intervention.

Table 10 shows the treatment effect for three outcomes: (1) the probability of enrolling in Principles of Macroeconomics one quarter after the intervention; (2) unconditional grade point average in Principles of Macroeconomics; and (3) grade point average in Principles of Macroeconomics conditional on enrolling. The difference between columns 2 and 3 is that column 2 includes the full sample of students, where the grade point average for students who *did not enroll in Principles of Macroeconomics* is coded to zero. Column 3 instead considers only the sample of students who enrolled in, and received a grade in, Principles of Macroeconomics. While Column 3 can not be explicitly interpreted as a "treatment effect" it may provide some insight regarding the presence of potential selection bias.

There are a few interesting takeaways from this table. First, Column 1 reveals that the treatment increases the likelihood of enrollment in Principles of Macroeconomics for URM students by 12.3 percentage points. Although there is no average treatment effect on course performance (likely due to attenuation of the estimate from the large number of zeroes in the outcome) Column 3 provides an interesting second result. Specifically, conditional on enrolling in Principles of Macroeconomics, URM students exposed to the treatment score around 0.51 GPA points lower than URM students in the control group. This difference is significant at the 5% level.

The presence of performance differences in the conditional sample despite no overall average treatment effect suggests that the treatment may be inducing *differential selection* into future Economics courses. Specifically, the results suggest that the URM students from the treatment group who select into Principles of Macroeconomics may be less prepared for the course, or lower performing than the URM students selecting in from the control group. While it is tempting to interpret this as a negative impact of the intervention, such selection is consistent with the expected effects of this intervention. The theory proposed in

Table 10: Treatment Effects on Principles of Macroeconomics Enrollment and Performance for URM Students One Quarter After the Intervention

	Full Sample		Conditional on Enrollment
	Prob. Enrolled	Grade Points	Grade Points
	(1)	(2)	(3)
Treatment x URM	0.147*	0.0894	-0.314
	(0.0824)	(0.228)	(0.242)
Treatment	-0.0243	-0.130	-0.196**
	(0.0370)	(0.118)	(0.0944)
URM	-0.0468	-0.0210	-0.107
	(0.0630)	(0.177)	(0.157)
Constant	0.320***	0.759***	2.781***
	(0.0565)	(0.175)	(0.134)
Demographic Controls	Yes	Yes	Yes
F-Stat for Interaction	2.763	0.0427	5.286
p-Value for Interaction	0.0969	0.836	0.0222
Observations	785	785	323
R-squared	0.159	0.186	0.475

Standard errors in parentheses. Observations are at the individual level.

All specifications include a full set of demographic controls and control for baseline academic performance.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

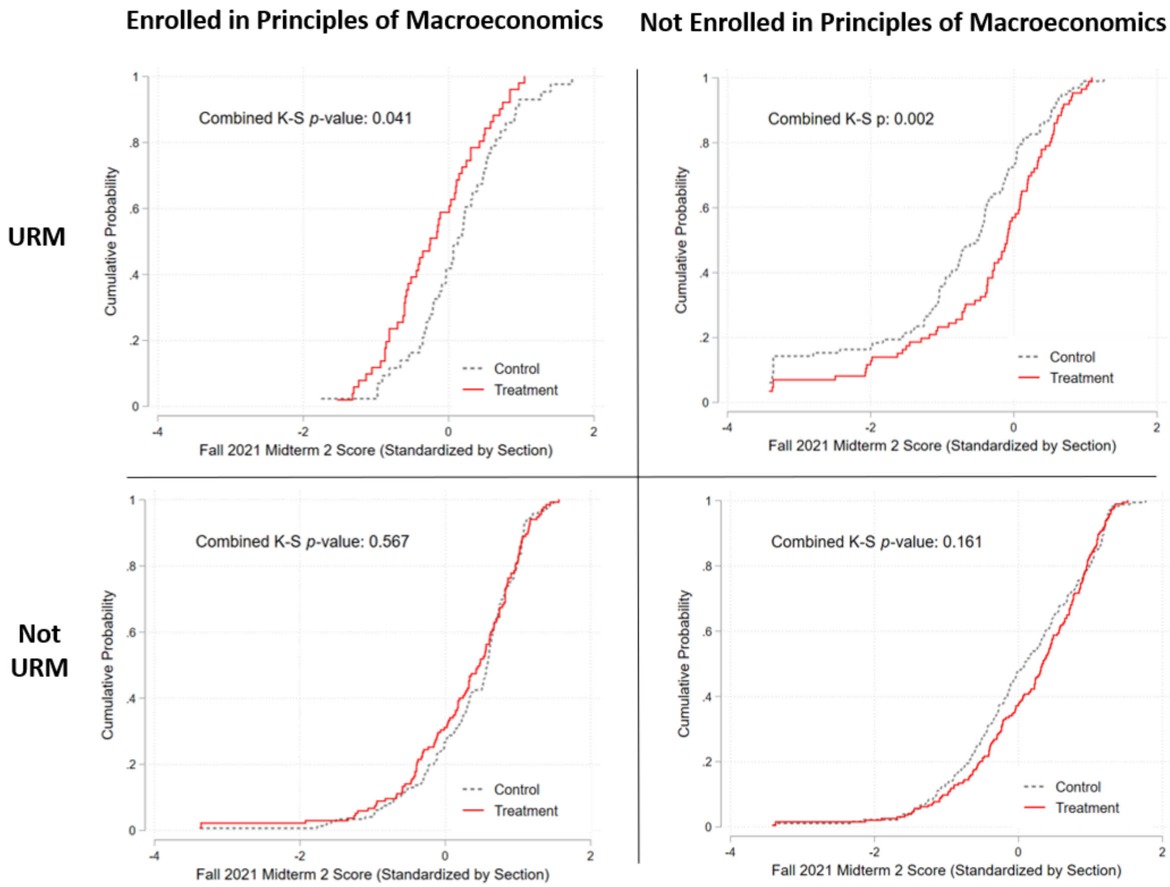
Section 2 suggests that information provision on a single major should only affect behavior for students who were not planning to pursue Economics at baseline. Of those students, it is reasonable that the *most likely group* to be impacted by the intervention is students on margin between pursuing and not pursuing Economics. Assuming that students with high grades and a clear aptitude for Economics will always enter Economics and students with the lowest grades and no aptitude for Economics will never enter Economics, the "compliers" in this study would be those "marginal" students with lower Economics performance and potentially low but positive baseline interest in Economics. If the information intervention increases the perceived value of majoring in Economics for these marginal students, this increase in interest may outweigh low baseline performance in their decision to enroll in a subsequent course.

To test whether the students selecting into a subsequent course from the treatment group are *lower-performing* than the students who enroll from the control group I examine cumulative distribution functions of grades in Principles of Microeconomics for subsamples of students who *enrolled* and *did not enroll* in the subsequent Principles of Macroeconomics course. If the URM students choosing to enroll in Principles of Macroeconomics from the treatment group are lower-performing than their URM counterparts in the control group, we would expect the distribution of their Fall course grades to be lower. Figure 2 shows the cumulative distribution functions for the treatment and control groups within subgroups defined by URM-status and enrollment one quarter after the intervention. For simplicity, I will use the term "Fall" to refer to outcomes in Principles of Microeconomics and "Winter" to refer to outcomes in Principles of Macroeconomics=.

A simple visual comparison between the top row and bottom row of this figure reveals a compelling pattern. For non-URM students, the Fall midterm grades for the treatment and control groups have highly similar distributions. This is true both for non-URM students who enrolled and did not enroll in the Winter course. This is what one would expect the graphs to look like if there was *no differential selection into Winter enrollment*. A two-sample Kolmogorov–Smirnov test reveals no stochastic dominance between the treatment and control grade distributions for either enrolled or non-enrolled non-URM students.

The top row, on the other hand, shows clear stochastic dominance. Specifically, for URM students who enrolled in Principles of Macroeconomics the distribution of Fall grades for the control group stochastically dominates the distribution of grades for the treatment group. This reveals that the URM students selecting into the next Economics course one quarter after the intervention are *lower-performing* than the URM students from the control group who enrolled. The Kolmogorov–Smirnov test reveals that this gap between the treatment

Figure 2: CDF Plots of Fall Midterm Grades for Treatment and Control Groups, by URM-Status and Enrollment in Principles of Macroeconomics One Quarter After Intervention



and control groups is significant at the 5% level.

This graph is consistent with the hypothesis that the conditionally lower performance of URM students who enrolled in a subsequent course is driven by differential selection. The CDFs provide compelling evidence that the information intervention induces lower-performing URM students to enroll in the next Economics course. These results are also consistent with one of the primary hypotheses that motivated this paper: that students who lack information about the field of Economics, may also lack the necessary background and preparation to initially succeed in the field. Consider, for example, a URM student on the margin between pursuing an Economics major and not pursuing the major. If they enter college with less prior knowledge about the field of Economics and lower preparation for a degree in Economics their performance in an introductory Economics course may reflect that. Absent information provision, their perceived utility from pursuing a degree in Economics may be low at baseline due to a lack of full information about the field. Given that low baseline expected utility, a low grade in an introductory course may nudge them to *leave the field*. The differential selection visible in this sample indicates that information provision may have the potential to overcome the negative impacts of a low introductory course grade. In Section 6 I explore the specific areas of information updating that may be driving this selection.

5.3 Effect on Majoring

The results thus far provide convincing evidence that a low-touch information intervention has significant impacts on URM course enrollment one quarter after the intervention. While this is a valuable outcome on its own, we may be more interested in the persistence of these short-term enrollment effects into long-term majoring decisions. If information provision does reduce a barrier to entry to Economics, we would expect that the treatment would affect students' likelihood of declaring an Economics major in the medium and long-run. If, on the other hand, there is no detectable average treatment effect on majoring, there are two potential explanations. The first is that the intervention is too low-touch to have long-run persistence. This is in line with what previous studies (Bayer et al. (2019)) have found. The other explanation is that while a lack of information may be one barrier to entry for URM students, it is not the primary barrier.

One reason this specific intervention may have greater long-run persistence than the Bayer et al. (2019) intervention is due to the population it is targeting. Bayer et al. (2019) run their information intervention prior to the start of the first semester for incoming college freshmen. An intervention targeting students who have not yet expressed interest in Economics addresses one of the two key blockages in the pipeline for students into the Economics major: *initial enrollment*. The intervention in this paper, on the other hand, targets students who have already expressed some interest in Economics (through their enrollment in a Principles of Microeconomics course). Thus in the context of the pipeline to an Economics major, this intervention focuses on the issue of *retention* after initial enrollment. This intervention may have greater success in affecting majoring behavior in part because it targets students at a later point in the majoring pipeline, and in part because it targets a sample of students that is more likely to have developed a preliminary preference ranking over majors that places Economics toward the top.

Currently, I am only able to examine majoring behavior in the medium-term (for up to two academic quarters following the intervention). Given that 78% of the study sample consists of first-year college students, it is possible that any major switching will not occur until over a year after the intervention (since declaring a major in the second year of college is fairly common at many universities). If this is the case, treatment effects on major switching may not emerge in the medium-term. Ongoing data collection will allow me to identify any longer-run effects. Table 11 shows no identifiable average treatment effect on the *probability of being an Economics major* for URM students within the academic year of the intervention. Unsurprisingly, the biggest predictor of being an Economics major in the medium-term, is entering college as an Economics major.

Table 11: Treatment Effect on the Likelihood of Being an Economics Major for URM Students

	One Term		Two Terms		Three Terms	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment x URM	0.0860	0.00339	0.0877	0.00800	0.101	0.0216
	(0.0885)	(0.0628)	(0.0894)	(0.0690)	(0.0884)	(0.0700)
Treatment	-0.0860**	-0.0244	-0.0738*	-0.0139	-0.0737*	-0.0173
	(0.0388)	(0.0265)	(0.0388)	(0.0283)	(0.0387)	(0.0300)
URM	-0.144**	-0.0948*	-0.124*	-0.0722	-0.145**	-0.0849
	(0.0627)	(0.0488)	(0.0632)	(0.0515)	(0.0622)	(0.0526)
Constant	0.491***	0.186***	0.485***	0.199***	0.479***	0.214***
	(0.0278)	(0.0434)	(0.0277)	(0.0459)	(0.0277)	(0.0501)
Demographic Controls	No	Yes	No	Yes	No	Yes
F-Stat for Interaction	2.92e-29	0.139	0.0297	0.00885	0.122	0.00469
p-Value for Interaction	1.000	0.710	0.863	0.925	0.727	0.945
Observations	801	785	801	785	801	785
R-squared	0.0122	0.557	0.00837	0.477	0.0101	0.419

Standard errors in parentheses. Observations are at the individual level

* p<0.1, ** p<0.05, *** p<0.01

6 Effects on Belief-Updating

The behavioral effects described in previous sections suggest that information provision has the potential to affect short-run Economics course enrollment decisions. The channel through which we would expect this information to impact actions is belief updating. This section examines belief updating on two primary information areas addressed by the information treatment: research topics in Economics and careers in Economics.[11]. The data on belief updating comes from baseline and endline surveys that students completed during the quarter of the intervention. The surveys contained a common set of beliefs questions on the topics covered by information intervention. For each topic, I use matching questions on the baseline and endline surveys to identify changes from the baseline to the endline. In all questions, the "correct answers" are determined based on the information provided in the information treatment.[12]

The surveys contained three primary types of beliefs questions: selection questions, ranking questions, and Likert scale questions. The survey questions eliciting beliefs on research topics and careers in Economics asked students to select multiple options from a list. For analysis of these types of questions, I separate the potential answers into three categories: *traditional*, *untraditional*, and *placebo*. In the main specifications, traditional answers are defined as answers students would be most likely to select based on common perceptions of the field of Economics. Untraditional answers are defined as answers that are related to information provided in the intervention. Finally, placebo answers are defined as answers that are neither expected to be common answers to the question, nor are linked to the information provided in the treatment.

The model for each of these regressions is as follows:

$$(EndlineBeliefs)_i = \beta Treatment_i + \phi(BaselineBeliefs)_i + \alpha X_i + \varepsilon_i \tag{1}$$

where X_i is a vector of demographic controls. This is the most flexible model for "belief updating" - a regression on endline beliefs that controls for baseline beliefs.

[11] While survey data is available for beliefs on potential income in Economics and diversity in Economics, the analysis for these measures is ongoing and therefore was omitted from this draft.

[12] In cases where the set of "correct answers" could be defined in multiple ways, I adhere to the strictest set of correct answers based on the information provided. I plan to run robustness checks using alternative definitions of "correctness".

6.1 Research Topics in Economics

The primary question on research topics in Economics asks students to select the topics they think they might "study in an Economics course". Table 12 shows the categorization of answers based on the previously defined categories. "Correct" topics include both "traditional" and "untraditional" topics.

Question	*Select the topics you think you might study in an Economics course?*
Traditional	Poverty and Homelessness, Income Gap
Untraditional	Civil Rights and Racial Discrimination, Gender Inequality, Health Care Access, Education Reform, Criminal Justice Reform, Immigration, Climate Change
Placebo	Overpopulation, Gun Violence

Table 12: Question and answer choices for beliefs about topics related to the field of Economics

Table 13 shows the average treatment effect on belief-updating regarding research topics for URM students. Columns 1 and 2 consider the the overall fraction of "correct" topics selected, while the remaining columns look at the fraction of topics selected by category.

Columns 1 and 2 reveal that the treatment increases the fraction of correct topics selected by URM students by around 9 percentage points, which is around a 30% increase from the control mean for URM students. Looking at the categorical breakdown, it is clear that this increase is driven by an increase in the fraction of untraditional topics selected, with no identifiable effects for traditional and placebo topics. This is the expected result given that the information treatment primarily draws attention to the untraditional topics, and therefore would be expected to shift beliefs on this dimension.

We may wonder whether this effect is driven by a lack of prior exposure to Economics (as was hypothesized in previous analyses) or by some other shared characteristic of URM students. One way to assess this is to look at heterogeneity by socieconomic status rather than URM-status. Table 14 looks at treatment effects for first generation students (defined as students with a highest parental education of less than a 4-year degree). The effects here are much lower in magnitude and insignificant. It is worth recalling that 55% of the sample of URM students identifies as first generation; however, only 39% of the sample of

Table 13: Average Treatment Effect on Fraction of Topics by URM Status

	Correct		Traditional		Untraditional		Placebo	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment x URM	0.130***	0.117**	0.0945	0.0874	0.245***	0.221**	0.0310	0.0245
	(0.0479)	(0.0479)	(0.0661)	(0.0660)	(0.0876)	(0.0875)	(0.0528)	(0.0538)
Treatment	-0.0288	-0.0251	-0.0130	-0.0112	-0.0575	-0.0499	0.00912	0.0105
	(0.0198)	(0.0205)	(0.0273)	(0.0280)	(0.0377)	(0.0387)	(0.0248)	(0.0254)
URM	-0.0772**	-0.0490	-0.0470	-0.0203	-0.147**	-0.0970	-0.0315	-0.0220
	(0.0332)	(0.0356)	(0.0489)	(0.0508)	(0.0601)	(0.0653)	(0.0368)	(0.0401)
Constant	0.228***	0.194***	0.564***	0.546***	0.321***	0.247***	0.248***	0.240***
	(0.0185)	(0.0336)	(0.0356)	(0.0535)	(0.0298)	(0.0610)	(0.0202)	(0.0378)
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-Stat for Interaction	5.411	4.521	1.830	1.610	5.634	4.730	0.737	0.542
p-Value for Interaction	0.0203	0.0338	0.176	0.205	0.0179	0.0300	0.391	0.462
Control URM Mean	0.292	0.292	0.560	0.560	0.377	0.377	0.282	0.282
Observations	801	785	801	785	801	785	801	785
R-squared	0.282	0.289	0.0715	0.0790	0.299	0.307	0.123	0.125

Standard errors in parentheses

Observations are at the individual level

Dependant variable is the endline fraction of topics selected

All specifications control for the rate at baseline

* p<0.1, ** p<0.05, *** p<0.01

Table 14: Average Treatment Effect on Fraction of Topics by First Generation Status

	Correct		Traditional		Untraditional		Placebo	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment x First-Gen	0.0201	0.0181	-0.00992	0.00294	0.0489	0.0390	-0.0257	-0.0294
	(0.0410)	(0.0420)	(0.0595)	(0.0601)	(0.0756)	(0.0778)	(0.0480)	(0.0499)
Treatment	-0.00829	-0.00625	0.0165	0.00566	-0.0257	-0.0158	0.0214	0.0255
	(0.0195)	(0.0214)	(0.0272)	(0.0289)	(0.0361)	(0.0403)	(0.0235)	(0.0260)
First Generation	-0.0464	-0.0408	-0.0236	-0.0514	-0.0832	-0.0707	0.0176	0.0225
	(0.0282)	(0.0304)	(0.0424)	(0.0442)	(0.0517)	(0.0563)	(0.0313)	(0.0341)
Constant	0.223***	0.208***	0.486***	0.580***	0.312***	0.268***	0.227***	0.219***
	(0.0170)	(0.0293)	(0.0300)	(0.0482)	(0.0276)	(0.0518)	(0.0188)	(0.0330)
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-Stat for Interaction	0.108	0.105	0.0157	0.0265	0.121	0.120	0.0106	0.00851
p-Value for Interaction	0.743	0.745	0.900	0.871	0.728	0.729	0.918	0.927
Control First-Gen Mean	0.367	0.367	0.706	0.706	0.473	0.473	0.358	0.358
Observations	937	793	937	793	937	793	937	793
R-squared	0.282	0.282	0.130	0.0727	0.288	0.301	0.133	0.128

Standard errors in parentheses. Observations are at the individual level.

Dependant variable is the endline fraction of topics selected. All specifications control for the rate at baseline.

* p<0.1, ** p<0.05, *** p<0.01

students identifying as first generation are also URM. The absence of identifiable treatment effects amongst this larger subsample indicates that there may be some characteristic of URM-identity that correlates with beliefs about research topics but is uncorrelated with socieconomic status. For example, it may be the case that URM students care more strongly about pursuing a major with the potential for impact on a broad range of social issues. This preference for"social impact" would not necessarily correlate with socioeconomic status, but may correlate with racial identity.

6.2 Careers in Economics

To identify beliefs about the types of careers they could pursue witha degree in Economics, students were asked to select the career sectors that they thought they would be "qualified to work in within 10 years of graduating with an undergraduate degree in Economics". Table 15 shows the exact question and the categorization of the answer set.

Question	*Select all of the career sectors that you think you would be qualified to work in within 10 years of graduating with an undergraduate degree in Economics? (Note: For sectors that may require a post-graduate degree, consider whether an undergraduate degree in Economics would qualify you to achieve such a degree.)*
Traditional	Finance, Business
Untraditional	Consulting, Government and Non-Profit, Academic or Policy Research, Law
Placebo	Medicine, Technology

Table 15: Question and categorization of answers for beliefs about careers in Economics

Tables 16 shows the treatment effect on knowledge about careers for URM students. Unlike the strong effect on research topics, this table reveals no meaningful belief updating on the dimension of careers in Economics. The fact that URM students seem to respond to information about research topics in Economics moreso than information about potential careers may be related to the fact that the information intervention has the largest behavioral impacts for lower-performing URM students. The student attracted by this intervention may be students with a *strong potential interest* in Economics research who were not previously aware that their research interests could be pursued with a degree in Economics. If those students had not previously intended to pursue Economics due to a lack of interest in the field, they may also have lower preparation to succeed in the field.

6.3 Income in Economics

To identify students' beliefs about the potential income attainable in Economics, students were asked to rank the a set of majors in order of "the median income you would expect someone with a bachelor's degree in that major to make in the U.S. 10 years after graduat-

Table 16: Average Treatment Effect on Fraction of Careers by URM Status

	Correct		Traditional		Untraditional		Placebo	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment x URM	-0.0128	-0.0154	0.0222	0.0181	-0.0209	-0.0249	-0.0171	-0.0167
	(0.0501)	(0.0503)	(0.0567)	(0.0552)	(0.0564)	(0.0576)	(0.0520)	(0.0534)
Treatment	0.0407*	0.0436*	0.0310	0.0355	0.0451*	0.0478*	0.0292	0.0274
	(0.0218)	(0.0224)	(0.0256)	(0.0261)	(0.0245)	(0.0250)	(0.0232)	(0.0237)
URM	0.0171	0.0202	0.0233	0.0205	0.00500	0.0135	-0.0111	0.00159
	(0.0374)	(0.0406)	(0.0442)	(0.0459)	(0.0413)	(0.0458)	(0.0384)	(0.0406)
Constant	0.380***	0.363***	0.676***	0.683***	0.311***	0.271***	0.133***	0.186***
	(0.0273)	(0.0419)	(0.0473)	(0.0599)	(0.0209)	(0.0415)	(0.0172)	(0.0371)
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
F-Stat for Interaction	0.383	0.394	1.109	1.217	0.225	0.195	0.0682	0.0504
p-Value for Interaction	0.536	0.531	0.293	0.270	0.635	0.659	0.794	0.822
Control URM Mean	0.481	0.481	0.665	0.665	0.389	0.389	0.151	0.151
Observations	801	785	801	785	801	785	801	785
R-squared	0.124	0.136	0.0305	0.0492	0.174	0.181	0.111	0.127

Standard errors in parentheses
Observations are at the individual level
Dependant variable is the endline fraction of careers selected
All specifications control for the rate at baseline
* p<0.1, ** p<0.05, *** p<0.01

ing". The majors included here were Electrical Engineering, Computer Science, Chemistry, Political Science, Biology, and Economics.

The outcome of interest here is the distance between the student's rank of Economics and the "true rank" which is defined based on the Bureau of Labor Statistics rankings. In the following table, the true rank is defined as 3 in a ranking from 1 to 6. Table 17 shows the results of these regressions. As is evident from both columns there is no treatment effect for URM students on their income ranking of Economics. It is possible that examining the raw distance between a student's ranking and the true ranking is too noisy of a measure to find an identifiable effect.

Table 17: Effect of Treatment on Income Rank Distance

	Rank Distance	
	(1)	(2)
Treatment x URM	-0.204	-0.186
	(0.192)	(0.192)
Treatment	0.0340	0.0341
	(0.0846)	(0.0853)
URM	-0.187	-0.0915
	(0.123)	(0.129)
Baseline Income Rank Dist.	0.565***	0.547***
	(0.0323)	(0.0337)
Constant	-0.187***	-0.231*
	(0.0631)	(0.128)
Demographic Controls	No	Yes
F-Stat for Interaction	0.971	0.768
p-Value for Interaction	0.325	0.381
Observations	720	709
R-squared	0.346	0.366

Standard errors in parentheses
Observations are at the individual level
* p<0.1, ** p<0.05, *** p<0.01

7 Discussion

Systemic racial gaps in higher education have been a focal point of policy discussions in the past few decades. This study contributes to a small but growing literature that attempts to better understand the causes of systemic racial gaps, particularly in high-return and high-impact fields such as economics. Specifically, I run a large randomized controlled trial to test the effect of an intervention designed to address information barriers that may adversely deter underrepresented minority (URM) students from pursuing the study of economics.

The results show that such a low-touch and highly scalable information intervention increases URM students' likelihood of enrolling in a subsequent Economics course by 12.3 percentage points.The magnitude of this effect is large given the low-touch nature of the intervention. Moreover, the information intervention seems to shift the composition of URM students who choose to enroll in the next economics course by drawing in lower-performing students. While this result might initially seem counterintuitive, it is consistent with the theory that in the presence of inaccurate or incomplete information about a field, information provision may shift a student's expected utility from majoring in that field by drawing their attention to aspects of the field that align with their interests and goals. This theory is corroborated by the evidence that URM students primarily update their beliefs about the research topics associated with Economics, rather than career-options and expected income. Absent information provision, these lower-performing students may have been deterred from enrolling in a subsequent course due only to their performance in an introductory course.

This work provides two valuable takeaways for education policies. First, this paper suggests that the structure of undergraduate economics programs should account for the possibility that students from different backgrounds may enter college with different levels of information about the field of economics. Bayer et al. (2019) shows that information about the scope and diversity of the economics field leads to a large increase in first-generation students' likelihood of enrolling in an economics course in their first semester. My study builds on this work by showing that information on expected income, career-options, scope of research topics, and diversity of researchers in economics can increase underrepresented minority students' likelihood of enrolling in a second economics course. The combination of these results suggests that information provision is valuable both for incoming students who have not yet expressed interest in economics, and for students who reveal a baseline interest in the field. Moreover, they suggest that students may care about receiving information on more than just the financial returns to a major. Economics programs may benefit from continuously providing students with information about the field at different stages of their undergraduate education. Furthermore, the large shifts in enrollment that result from a low-touch email intervention suggests that more intensive information interventions could have even greater impacts on retention of URM students.

A second takeaway for education policy is that while information provision seems to have short-run impacts on URM students' persistence in economics, the current evidence suggests that this may not translate into long-run persistence in the economics major. I find no effect on my intervention on the likelihood of declaring an Economics major. Similarly, Bayer et al., find no effect that the effects of their intervention persist beyond initial course enrollment. While the lack of long-run effects may be due to the low-touch nature of these interventions, it may also indicate that there are barriers to persistence beyond just incomplete information or stereotypes. For example, my paper shows that the information intervention induces *lower-performing* URM students to enroll in a subsequent course; however I also find that conditional on enrolling in a next course, students in the treatment group perform worse than students who enrolled from the control group. This suggests that while full information may increase the incentives for students to stay in the economics major, if these students lack the academic preparation to perform well in the major they may be deterred from the field in the long run. One way to address this is to ensure that there are additional academic supports for students who are struggling in introductory courses.

These results propose several avenues for future research. First, while the low-touch and no-cost information provision in this study has the benefit of being easily salable, the evidence suggests that it may not lead to persistent long-run effects. Future work should explore the effect of more high-touch information interventions on similar outcomes, in order to assess whether the relative gains from increasing the "intensity" of the intervention are commensurate to the added costs of running and scaling those interventions. Secondly, this study suggests that URM students may care about features of majors beyond just expected income; however, little is known about exactly what features of college majors influence their decisions. Identifying the factors that are most valuable for decision-making will help future researchers and policymakers develop well-targeted information interventions to address barriers to entry in economics, and STEM fields more broadly. Finally, there is little evidence about the role that educational and academic resources play in increasing long-run retention of minority students in rigorous, technical fields. Research on barriers to entry and persistence beyond incomplete information is critical to understanding which barriers are strongest for underrepresented minority students, and what combination of interventions is necessary to achieve fully equitable access to these high-return majors.

References

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

A. Bayer, S. P. Bhanot, and F. Lozano. Does simple information provision lead to more diverse classrooms? evidence from a field experiment on undergraduate economics. 109: 110–14, 2019.

M. Beffy, D. Fougere, and A. Maurel. Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics*, 94(1):334–347, 2012.

M. Carlana. Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224, 2019.

S. E. Carrell, M. E. Page, and J. E. West. Sex and science: How professor gender perpetuates the gender gap. *The Quarterly journal of economics*, 125(3):1101–1144, 2010.

A. Chambers, S. Dickert-Conlin, C. Elder, S. J. Haider, and S. Imberman. Info. econ: Increasing diversity among economics majors. In *AEA Papers and Proceedings*, volume 111, pages 133–37, 2021.

CollegeBoard. Technical report, College Board, 2019. URL https://reports.collegeboard.org/media/pdf/Program-Summary-Report-2019_1.pdf.

J. De Quidt, L. Vesterlund, and A. J. Wilson. Experimenter demand effects. In *Handbook of research methods and applications in experimental economics*. Edward Elgar Publishing, 2019.

R. W. Fairlie, F. Hoffmann, and P. Oreopoulos. A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, 104(8): 2567–2591, 2014.

I. Haaland, C. Roth, and J. Wohlfart. Designing information provision experiments. 2020.

G. Hale and T. Regev. Gender ratios at top phd programs in economics. *Economics of Education Review*, 41:55–70, 2014.

G. A. Hoover and E. Washington. Report: Committee on the status of minority groups in the economics profession (csmgep). In *AEA Papers and Proceedings*, volume 111, pages 764–79, 2021.

M. S. Kofoed et al. The effect of same-gender or same-race role models on occupation choice: evidence from randomly assigned mentors at west point. *Journal of Human Resources*, 54 (2):430–467, 2019.

C. Porter and D. Serra. Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–54, 2020.

M. Wiswall and B. Zafar. Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824, 2015.

A Tables and Figures

Table A.1: Balance of covariates for course Section B

	Control	Treatment	t-test(p)
Female	0.401	0.397	0.94
URM	0.217	0.171	0.33
International	0.387	0.319	0.23
Bins of Parental Educ.	1.167	1.271	0.28

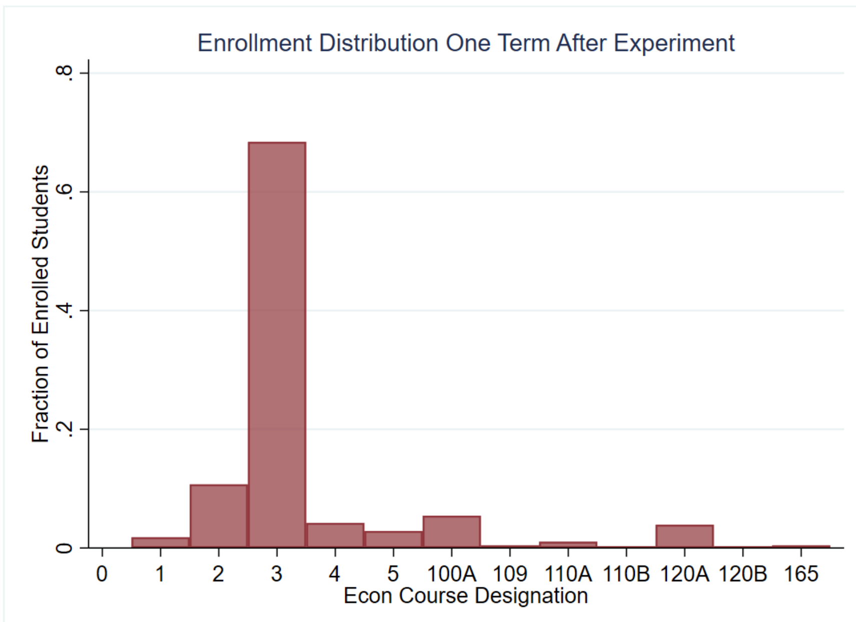


Figure A.1: Distribution of course enrollments one term after the intervention.

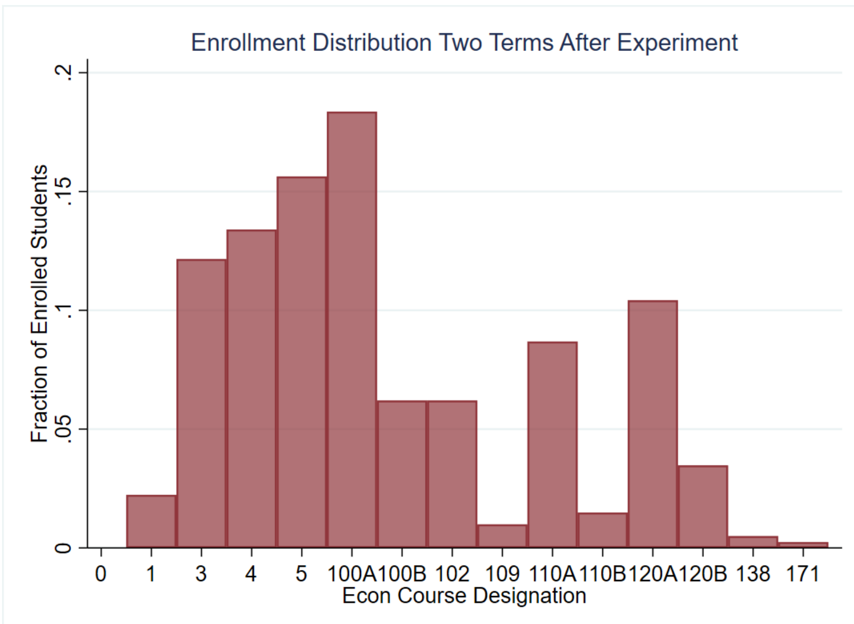


Figure A.2: Distribution of course enrollments two terms after the intervention.

B Survey Materials

Dear Students,

I would like to officially welcome you to the Economics Department at UCSD! I hope your first quarter is going smoothly. As you begin your journey in Economics, I encourage you to look through some of the resources in this email and the attached flyer.

Student Groups:

- Women And Minorities in Economics (WAMIE): https://ucsdwamie.ucsd.edu/
- Undergraduate Economics Society (UES): https://ues.ucsd.edu/
- UCSD Women in Business: https://www.ucsdwib.com/

Department Resources:

- Department Website: https://economics.ucsd.edu/undergraduate-program/index.html
- EconUG Blog: https://econugblog.wordpress.com/

Figure B.1: Common introduction for treatment and control emails

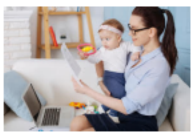
What is Economics?

MYTH: Economics is only about money.

FACT: Economics has applications across many disciplines, including **health**, **gender**, **the environment**, **education**, and **immigration**.

Economics can help answer questions such as:

How common is racial discrimination and profiling during traffic stops?



Why do mothers earn less money?

Do selective high schools help students from disadvantaged backgrounds get into better colleges?

What careers are possible with a degree in Economics?

MYTH: The Economics major only prepares you for careers in Finance or Business.

FACT: An undergraduate degree in Economics can prepare you for a wide range of careers in areas such as:

- **Policy research** at a Think Tank or a Non-Government Organization (NGO)
- **Government** or **non-profit research**
- **Law**
- **Economic consulting** on issues of healthcare, anti-trust, energy and the environment, and insurance.
- Pursuing a **Ph.D. in Economics**

Figure B.2: First half of the body of the treatment email

How much money can you earn with a degree in Economics?

MYTH: Economics majors make less money than Engineering majors

FACT: The median salary for Economics majors with around 10 years of work experience is around $140,000 and is comparable to the median salaries in fields such as Engineering and Computer Science.

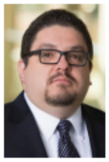
Meet Some Economists

M. V. Lee Badgett Professor of Economics at UMass - Amherst

"I study LGBT inequality to learn how to end it."

Cecilia E. Rouse Current Chair of the Council of Economic Adviser

"I [was] drawn to study… the reasons that jobs disappear; the impact of education on people's job prospects…"



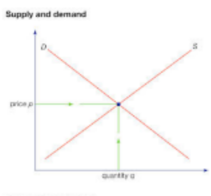
Mark Hugo Lopez Director of Hispanic Research, Pew Research Center

"The questions I was interested in – the economic life of U.S. Latinos, immigration… – were topics I could study with a degree in economics."

Figure B.3: Second half of the body of the treatment email

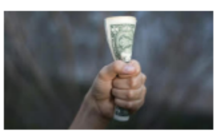
What is Econ 1?

Econ 1 is designed to help you understand why individuals interact in market settings. You will learn:



How the demand for, and supply of, a product is determined and how equilibrium in a market occurs.

How individuals make consumption choices and why perfectly competitive markets lead to efficient outcomes.



The effects of the minimum wage and the factors contributing to income inequality

International trade and how to determine the winners and losers from trade.

Figure B.4: Body of the control group email

Humanities: Linguistics, Literature, History, Music, Philosophy, Theatre, Visual Arts, Other Arts

Education and Communications: Education Sciences, Communication

Economics: Economics, Business Economics, Management Science, Mathematics-Economics (Joint degree)

Politics: Political Science, International Studies

Math & Physics: Mathematics, Probability and Statistics, Physics

Biology and Chemistry: General Biology, Microbiology, Neurobiology, Chemistry, Biochemistry, Cognitive Science

Earth Sciences: Geoscience, Marine Biology, Oceanic and Atmospheric Sciences

Social Science: Anthropology, Psychology, Sociology, Ethnic Studies, Other Social Sciences

Engineering & Computer Science: Bioengineering, Electrical, Aerospace, Mechanical, Chemical, Structural, Computer Engineering, Computer Science

Figure B.5: The list of "major groups" students were shown throughout the survey. This approach was used to reduce cognitive load for survey respondents and increase ease of analysis.

C Theoretical Model

C.1 Setup: A Model with Two Majors

Consider a simplified model of a student's choice between two majors. Assume that if the student had perfect information about both majors, they would have a strict preference ranking over the two majors. Note that this is akin to the assumption that the true utilities from the two majors differ by some amount $\varepsilon > 0$. Define the student's major choice m as

$$m \in \{m_1, m_2\}$$

where the utility from major m is a function of long-run income (I_m) and long-run satisfaction (S_m).[13]. We can define a student's expected utility from major m as

$$U_m = F(\mathbb{E}[I_m], \mathbb{E}[S_m])$$

where $\mathbb{E}[I_m]$ and $\mathbb{E}[S_m]$ are the student's expectations of the level of income and satisfaction they will achieve by majoring in m. Define the true utility of major m as

$$U_m^* = F(I_m^*, S_m^*)$$

where I_m^* and S_m^* are the *true* levels of income and satisfaction the student would attain from pursuing major m. Note that if a student has perfect information about major m, $\mathbb{E}[I_m] = I_m^*$, $\mathbb{E}[S_m] = S_m^*$, and $U_m = U_m^*$.

C.2 Consequential and Inconsequential Information

In the above two-major setting, we can define an *information gap for major m* as occurring when:

$$U_m \neq U_m^* \tag{2}$$

In order to assess the potential effects of information provision in the presence of such information gaps, consider the following ratio:

$$\frac{U_{m_2} - U_{m_1}}{U_{m_2}^* - U_{m_1}^*} \tag{3}$$

[13]Note that satisfaction is a broad term that can include satisfaction attained from jobs that result from majoring in m as well as the satisfaction gained from the major itself (i.e. coursework, community, passion for the subject)

The numerator in 3 provides information about the student's preference ranking over majors in the presence of an information gap, while the denominator provides information about the student's preference ranking under perfect information. Therefore this ratio can be thought of as an indicator of how *optimal* a student's decision is. Note that if the student has perfect information, $U_{m_1} = U^*_{m_1}$ and $U_{m_2} = U^*_{m_2}$, resulting in

$$\frac{U_{m_2} - U_{m_1}}{U^*_{m_2} - U^*_{m_1}} = 1$$

In the two-major case, a student is making a sub-optimal decision if their expected utilities of the two majors lead to a preference ranking that is different than their preference ranking under true utilities. In other words, a sub-optimal major choice occurs when

$$\frac{U_{m_2} - U_{m_1}}{U^*_{m_2} - U^*_{m_1}} < 0 \tag{4}$$

and an optimal major-choice occurs when

$$\frac{U_{m_2} - U_{m_1}}{U^*_{m_2} - U^*_{m_1}} > 0 \tag{5}$$

Now, consider the case where information is provided on either one or both majors. In order to establish a framework for understanding how this information will shift behavior, I first define two types of information provision: *inconsequential* and *consequential*. *Inconsequential information provision* can be defined as information provision that does not affect the student's choice of major. This can be represented as an *unchanged sign* of the information gap. Take for example, the case where

$$\frac{U_{m_2} - U_{m_1}}{U^*_{m_2} - U^*_{m_1}} > 0$$

and after information provision on one or both majors

$$\frac{U'_{m_2} - U'_{m_1}}{U^*_{m_2} - U^*_{m_1}} > 0$$

where U'_{m_2} and U'_{m_1} are the updated expected utilities. In this example, the information does not change the sign of the information gap. In order for this to be the case, it must be true that

$$\frac{U'_{m_2} - U'_{m_1}}{U_{m_2} - U_{m_1}} > 0 \tag{6}$$

Note that inconsequential information provision can also occur under initially sub-optimal preferences, so long as Equation 6 holds.

The alternative is *consequential information provision* which can be defined as information provision that leads the student to change their major choice. Note that again, consequential information provision can occur under both optimal and sub-optimal initial preferences so long as the condition in Equation 7 holds.

$$\frac{U'_{m_2} - U'_{m_1}}{U_{m_2} - U_{m_1}} < 0 \tag{7}$$

The remainder of this section will focus on consequential information provision, which captures the critical case where information not only affects expected utilities, but actions as well. In the simple two-major example, if the student is making a *sub-optimal* decision at the baseline, consequential information provision will necessarily be utility-improving (since it must move the student from the sub-optimal choice to the optimal choice).

Conversely, if the student is making the optimal decision at the baseline, consequential information provision must necessarily result in a decrease in utility. Specifically, if

$$\frac{U_{m_2} - U_{m_1}}{U^*_{m_2} - U^*_{m_1}} > 0$$

following the logic of Equation 7, consequential information will lead to

$$\frac{U'_{m_2} - U'_{m_1}}{U^*_{m_2} - U^*_{m_1}} < 0$$

Note that information provision in this model can apply to either m_1 or m_2. In each instance, there are a set of conditions that must hold in order for the information provided to be consequential. For simplicity, assume without loss of generality that:

$$U^*_{m_1} < U^*_{m_2} U_{m_1} < U_{m_2}$$

In other words, assume that at the baseline (with the presence of information gaps) the student's **preferred choice** is m_2, which is also the choice that will give the student the highest true utility.

Assume for further simplification, that the information provided to the student is "good information" which can be defined as information that shifts the student's expected utility in the direction of the true utility.

Consider first, information provided on m_1, the less preferred major at baseline. If the expected utilities of both majors are *underestimated*, we get the following inequalities:

$$U^*_{m_1} < U^*_{m_2}$$
$$U_{m_1} < U_{m_2}$$
$$U_{m_1} < U^*_{m_1}$$
$$U_{m_2} < U^*_{m_2}$$

Combining, we get two possible relationships between the true utilities, and the expected utilities at baseline:

1. $U_{m_1} < U_{m_2} < U^*_{m_1} < U^*_{m_2}$
2. $U_{m_1} < U^*_{m_1} < U_{m_2} < U^*_{m_2}$

In this context, if m_2 is sufficiently underestimated (i.e. $U_{m_2} < U^*_{m_1}$), information on m_1 that causes the expected utility of m_1 to be updated to the true value, will be consequential.

Now consider information provided on m_2, the preferred major at baseline. If the expected utilities from both majors are *overestimated* we get the following inequalities:

$$U^*_{m_1} < U^*_{m_2}$$
$$U_{m_1} < U_{m_2}$$
$$U_{m_1} > U^*_{m_1}$$
$$U_{m_2} > U^*_{m_2}$$

Combining, we get two possible relationships:

1. $U^*_{m_1} < U_{m_1} < U^*_{m_2} < U_{m_2}$
2. $U^*_{m_1} < U^*_{m_2} < U_{m_1} < U_{m_2}$

In this context, if m_1 is sufficiently overestimated (i.e. $U^*_{m_2} < U_{m_1}$), then information on m_2 that moves the expected utility to the perfect-information utility will be consequential.

C.3 Generalized Model over N majors

Consider an undergraduate student who is selecting a major, m from a set of N majors: $m \in \{m_1, m_2, \ldots, m_N\}$ WLOG let:

$$m_N = max_{m_n} \{U_{m_1}, U_{m_2}, \ldots, U_{m_N}\}$$

we can define consequential information provision on some major m_j as occurring if:

$$\frac{U'_{m_N} - U'_{m_j}}{U_{m_N} - U_{m_j}} < 0 \tag{8}$$

Note that whether this information provision is utility-increasing or utility-decreasing depends on the sign of $U^*_{m_N} - U^*_{m_j}$. Consequential information provision will be *utility-increasing* if

$$\frac{U_{m_N} - U_{m_j}}{U^*_{m_N} - U^*_{m_j}} < 0$$

which occurs when $U^*_{m_N} - U^*_{m_j} < 0$ (since by construction, $U_{m_N} - U_{m_j} > 0 \ \ \forall j \neq N$).

The intuitive interpretation here is that if information is provided on a major that has a higher true utility than the student's current choice, the student will be better-off from switching to that major.

Conversely, information provision will be *utility-decreasing* if

$$\frac{U_{m_N} - U_{m_j}}{U^*_{m_N} - U^*_{m_j}} > 0$$

which occurs when $U^*_{m_N} - U^*_{m_j} < 0$. In other words, if information is provided on a major that has a lower true utility than their current choice, the student will be worse-off from switching to that major.