

# The Hidden Curriculum

Michael G. Cuna\*                      Elaine Shen  
University of Chicago                  UC Berkeley

**JMP**

October 31, 2025

*Please click here for the latest version.*

## **Abstract**

Despite dramatically expanded access to selective U.S. colleges, first-generation students persistently trail continuing-generation peers in GPA, internship attainment, and early-career outcomes. We identify a key mechanism: the hidden curriculum—unwritten strategies like cold-emailing alumni or strategically engaging faculty—essential for success yet unknown and costly to discover without guidance. Leveraging survey and administrative data from 100,000+ undergraduates across 20 public universities, we document stark disparities: first-generation students invest 14-26% less in these high-return hidden actions while over-investing in formal tasks. Standard explanations—income, ability, or preferences—explain little. Through a field experiment at UC Berkeley, we isolate causal channels by randomizing information on action availability (awareness) versus returns (beliefs): awareness treatments close the 30% baseline gap almost entirely. We then use an online AI-advisor experiment to expose underlying search frictions; first-generation students allocate just 11% (versus 16%) of queries to hidden topics and follow up 47% less. However an “active” AI that expands consideration sets narrows search gaps and reverses the follow up rates. By formalizing the hidden curriculum as dual informational frictions, we demonstrate that overcoming these invisible barriers requires more than equal access.

---

\*Cuna (michaeltcuna@uchicago.edu); Shen (elaineshen@berkeley.edu).

Michael Cuna is especially indebted to his advisors Leonardo Bursztyn, John List, and Alex Imas for their invaluable feedback. We thank Ned Augenblick, Alec Brandon, Brent Hickman, Justin Holz, Gautam Rao, Andrew Simon, for helpful comments. We are also grateful to Igor Chirikov for facilitating access to the SERU Consortium Data and for providing context. IRB of record: University of Chicago Protocol #IRB24-1552; UC Berkeley Protocol #2024-09-17838 (Experiment); UC Berkeley Protocol #2025-07-18799 (UCES Data); AEA Trial Registration #AEARCTR-0014970; AEA Trial Registration #AEARCTR-0016671.

# 1 Introduction

Across schools and firms, formal access to opportunity has expanded, yet sizable performance gaps persist among people who look similar on paper. A common feature of these expansions is that they make many individuals “firsts”: first in their family to enter a selective university, first to work in a professional office, first to navigate organizations whose rules are unfamiliar. Being first changes not only resources but also the informational environment. Newcomers face uncertainty surrounding the institutions that govern these organizations.

We argue that a key driver of these gaps is the hidden curriculum: the unwritten, context-specific norms and expectations that govern success within an organization. Unlike formal requirements, the hidden curriculum is unevenly known and costly to learn by trial and error. In our setting –undergraduate students at U.S. universities– this includes practices such as engaging with professors, networking, and participating in extracurricular activities such as professional student organizations. We hypothesize that limited knowledge of these practices, and subjective beliefs about their returns, depress investment among newcomers.

We study the hidden curriculum in career and academic choices of first-generation college students. Our evidence combines large-scale administrative and survey data with experiments that vary information about “what to do” and “what it’s worth,” allowing us to quantify the role of hidden curriculum information in students’ effort choices. We show that the hidden curriculum can explain differences in actions across first-generation and continuing-generation students. We also show how lack of knowledge about the hidden curriculum may exacerbate inequality by preventing already disadvantaged students from being able to learn actions and strategies crucial for academic and professional success.

In this paper, we show that beyond formal qualifications and skills, access to tacit knowledge and culturally conditioned behaviors significantly affects students’ post enrollment experiences. We start by developing a simple conceptual framework to give a clear understanding of the mechanisms of the hidden curriculum, and of the consequences for those affected by it. We start with a traditional skill-augmenting investment setting (Cunha et al. (2013); Cunha et al. (2022)). We then model the hidden curriculum as two distinct informational frictions: i) awareness of the existence of certain investments/actions, and ii) traditional subjective beliefs on returns of such actions. The model predicts that individuals who are impacted by either of the previous two margins will: i) invest less in the

hidden curriculum; ii) will invest more in the formal curriculum<sup>1</sup>. These two mechanisms also help us think through potential learning dynamics and policy relevant solutions.

For our empirical strategy we combine evidence from three empirical approaches to document the existence of the hidden curriculum. First, we provide evidence of differences in outcomes and actions between first-generation students and continuing-generation students, and how these differences persist over their careers. We use large observational and survey data collected by the Student Experience in the Research University Consortium based at the Center for Studies in Higher Education. The data for the 2022 academic year include more than 100,000 students from 20 public colleges across the US. We show that first-generation students: i) report worse outcomes in both GPA and internships; ii) report lower investments in actions that are part of the hidden curriculum (i.e. the actions that do not carry over from high school and that are not in the formal syllabus); iii) report higher investments in actions that are part of the formal curriculum; and iv) report small degrees of convergence of investments in actions throughout their college careers. We also use our rich dataset to show that both constraints (mainly income, baseline ability, language barrier, nationality) and preferences (willingness to engage in club and extracurricular activities) do not explain most of the variation in outcomes.

Second, since the observational data alone cannot determine the causal impact of information on actions, we control for the assignment mechanism by running a pre-registered field experiment in a large U.S. public university (UC Berkeley) where we recruit 645 students and elicit whether they are willing to pay for a service that provides help in an action that we consider part of the hidden curriculum: cold-networking. We then randomly assign students to either a control group or one of two information interventions to tease out one of the two mechanisms introduced in our model. The first information intervention is designed to just give students the knowledge of cold networking being an available strategy. The second information intervention is designed to update their beliefs on the returns of such actions. This design will answer whether: i) there is indeed an issue of hidden curriculum, by looking at the control group vs the treatment groups; and ii) determine if the issue is knowledge about awareness of actions or about returns.

The results show that in the control group first-generation students are 30% less likely to be willing to pay for the service. This is in line with our observational data on students' willingness to take actions in the hidden curriculum. Once we look into the effect of the treatments we see that lack of awareness of cold-networking as an available strategy

---

<sup>1</sup>As we will see in Section 2 this is true for most but not all cases.

explains most of the baseline gap in willingness to pay for the service. First-generation and continuing-generation students are statistically indistinguishable after the information treatment on availability is provided. We also find that more heavy handed information interventions that highlight the necessity of such strategies have a positive but not statistically significant effect. We hypothesize that the treatment can backfire if the student was completely unaware. Robustness checks show that the results are not sensitive to confounders that would usually be deemed as the major culprits of this differential in behavior across first-generation and continuing-generation students.

In addition, we show that information also plays a role in other groups that may not have the same insider knowledge about the institution and does not play a role for traditionally disadvantaged groups that face constraints other than information (e.g. gender). Regarding the former, we show how information treatment on returns changes the demand for the service of community college transfer students. We provide an explanation of this heterogeneity in treatment response in Section 4.2. The existence of heterogeneity in mechanisms across groups implies the need for a more scalable solution compared to traditional interventions or one-off nudges to increase access to on-campus resources.

Finally, to uncover mechanisms of learning, and to help propose a flexible tool that can help in both increasing awareness and beliefs on returns, we develop an experiment to look more in depth into what keeps the hidden curriculum hidden. In particular, we focus on students' effort in information search and their consideration sets. In a pre-registered online experiment we provide students with an AI college advisor tailored to help them with questions regarding behaviors and strategies that they can adopt both within and outside of college to succeed in the labor market. To uncover mechanisms we randomize students into one of two conditions: i) a Passive AI that only discusses topics directly related to the students question, and ii) an Active AI that purposefully informs students of hidden curriculum strategies in relation to the question they asked.

The results reveal striking differences in how first-generation and continuing-generation students approach information acquisition. At baseline, first-generation students devote only 10% of their queries to hidden curriculum topics compared to 16% for continuing-generation students. More importantly, their search patterns differ qualitatively: first-generation students ask more general questions about GPA requirements and whether internships matter, while continuing-generation students seek specific guidance on networking strategies and interview preparation. First-generation students also express greater uncertainty, often questioning whether strategies like networking are “really necessary” for

their major, and are far less likely to follow up when the AI introduces hidden curriculum topics (26% follow-up rate versus 49% for continuing-generation students).

The active AI treatment substantially increased engagement with hidden curriculum topics for both groups, effectively bringing first-generation students to similar active search rates for hidden curriculum topics as their peers<sup>2</sup>. However, qualitative analysis reveals that first-generation students face higher costs in processing this new information and are more skeptical regarding the quality of the information. These patterns help explain why information gaps persist in equilibrium: those who most need knowledge about informal strategies are least equipped to search for, recognize, and act on this information when they encounter it. The experiment uncovers that the hidden curriculum is not just about missing information, but about systematic differences in how students navigate the complex, ambiguous process of discovering what they need to know.

This paper contributes to several strands of literature. First, we bridge the sociology and economics literature. We do so by formally introducing the concept of the hidden curriculum (Jackson (1968)) into economics, and by using tools from applied microeconomics to show causal evidence of its existence. In sociology, the hidden curriculum are tacit rules, implicit expectations, and unwritten norms transmitted by institutions, and is distinct from the formal curricula (Portelli (1993); Hafferty (1998); Lawrence et al. (2018)). These frameworks, and adjacent work on cultural and social capital, explain why “knowing how” to navigate elite organizations (e.g., approaching faculty, initiating networking, reading situational norms) is unevenly distributed and consequential (Bourdieu (1986); Jack (2019)). These ideas are connected to empirical hiring research showing that cultural and social class matching shape who advances in elite labor markets (Rivera (2016); Granovetter (1973)).

On the economics side, our focus on labor market success leverages a tradition that models information frictions and networks (Stigler (1962); McCall (1970); Ioannides and Loury (2004); Calvo-Armengol and Jackson (2004); Topa (2011)) and recognizes that early-career search conditions have persistent effects on careers, underscoring the stakes of the transition we study (Kahn (2010); Oyer (2006); Oyer (2008)). Methodologically, we link the hidden-curriculum lens to formal models of limited consideration/attention (Masatlioglu et al. (2012); Manzini and Mariotti (2014)) and beliefs about returns and norms (Jensen (2010); Wiswall and Zafar (2015); Dizon-Ross (2019); Ba et al. (2024); Bursztyn

---

<sup>2</sup>Active search rate is defined as students own effort in searching information about hidden curriculum topics.

et al. (2020)), consistent with broader evidence on behavioral inattention and decision-making under imperfect information (DellaVigna (2009); Gabaix (2019)). Finally, our setting complements field experiments in job-search guidance that uncover how information and advice change search behavior and applications (Altmann et al. (2018); Belot et al. (2019)), and connects to recent evidence that informal interactions inside firms (“schmoozing”) affect advancement—one manifestation of tacit organizational norms (Cullen and Perez-Truglia (2022)).

Second, we advance the education literature along two dimensions. First, we propose and test a mechanism for persistent achievement gaps between first-generation and continuing-generation students that complements recent evidence documenting these disparities across academic and early-career outcomes (Aucejo et al. (2025); Karna et al. (2025); Stansbury and Rodriguez (n.d.)). While existing explanations emphasize pre-college preparation and resources or academic match (Bowen et al. (2009); Hoxby and Avery (2012)), we highlight a complementary channel: differential knowledge of informal strategies even among students at the same institution with similar observable credentials. Second, our experiments contribute to the literature on student beliefs and learning by showing that information acquisition itself is endogenous to initial information endowments: students who begin college less informed are systematically less likely to discover high-leverage, tacit strategies on their own. This connects to work documenting belief formation and updating about returns in education and effort (Jensen (2010); Zafar (2011); Wiswall and Zafar (2015); Arcidiacono et al. (2020); Rury and Carrell (2023); Ersoy (2023)), and to new evidence that first-generation students enter with greater uncertainty and knowledge gaps that sustain performance differences (Aucejo et al. (2025)).

Third, we use AI chat interfaces as a method to study information acquisition and learning, contributing to the empirical literature on human-AI interaction. Prior work shows that algorithmic advice and AI assistance systematically shape human decisions (Bundorf et al. (2019); Lai et al. (2021); Bastani et al. (2024); Jabarian and Henkel (2025)). Unlike traditional lab tasks or one-shot surveys, AI chat logs capture multi-turn conversations, clarifications, and reformulations, enabling direct and natural measurements of information search efforts (Mo et al. (2024); Paruchuri et al. (2025)). Moreover, AI agents can provide a less judgmental environment that increases self-disclosure on sensitive topics, thus reducing social desirability bias when participants consent to logging (Lucas et al. (2014); Croes and Antheunis (2024)). Our approach builds on two strands of the literature: (i) AI tutoring and education that infers learning dynamics from fine-grained interaction traces (Koedinger et al. (2015); Graesser et al. (2004); Koedinger

and Aleven (2007); Roll et al. (2007)), and (ii) the new availability of large-scale LLM conversational datasets that reveal real-world information-seeking patterns (Paruchuri et al. (2025); Chatterji et al. (2025)). Finally, because the same tool can deliver guidance while recording behavior, AI can serve both as research instrument and policy tool. Recent randomized and field evidence shows well-designed AI tutors improve learning and feedback provision (Kestin et al. (2025); Meyer et al. (2024)). While, other work cautions that unconstrained assistance can harm learning without guardrails (Bastani et al. (2024)), or that AI can increase disparities (Caplin et al. (2024)). Additionally AI can be used to complement quantitative findings with qualitative data (Chopra and Haaland (2023)). Similarly, we find qualitative differences in how first-generation and continuing-generation students interact with AI. First-generation students ask less precise questions and are less likely to follow up on AI nudges at baseline.

Finally, we contribute to research on inequality and social mobility by proposing a mechanism of why equal access alone rarely delivers equal opportunity. Policies that remove formal barriers—expanding college access, adopting need-blind admissions, and offering generous financial aid have increased enrollment at selective institutions, yet intergenerational mobility remains limited and highly uneven across colleges (Chetty et al. (2017); Chetty et al. (2020); Dynarski and Scott-Clayton (2013) ). Our evidence shows that even when first-generation students enter the same campuses as their more advantaged peers, a second, invisible barrier persists: success requires mastery of unwritten rules transmitted through family and social networks but seldom taught explicitly. This mechanism complements classic accounts emphasizing financial and human capital transmission across generations (Becker and Tomes (1979); Solon (1999)) and aligns with sociological evidence on cultural capital and classed know-how (Lareau (2002); Calarco (2011); Jack (2019)). With this we propose a mechanism to reconcile two facts: U.S. colleges have diversified their student bodies, yet the distribution of mobility contributions remains skewed and overall mobility rates stubbornly low. Our results suggest that equality of opportunity requires not only opening doors, but also equalizing access to the unwritten rules of the rooms students enter.

The paper proceeds as follows. In Section 2 we introduce a simple model of the hidden curriculum. In Section 3 we present motivating evidence documenting the hidden curriculum in a large survey of over 100,000 students from 20 public universities in the US. In Section 4 we show causal evidence through a field experiment of information frictions leading to under-utilizing hidden curriculum strategies. In section 5 we describe information search efforts and the role of consideration set nudges in learning the hidden curriculum.

Finally, we conclude in Section 6 and discuss the policy relevance and potential areas for further research.

## 2 Setting and Basic Framework

In this section we will formally define the hidden curriculum and introduce a simple framework. The framework will be helpful in understanding: i) patterns in the data, ii) mechanisms that define this information friction, and iii) policy implications.

### 2.1 Defining the Hidden Curriculum

We define the hidden curriculum as all those actions, strategies and expectations that are not formally required or stated by the institution. In the case of first-generation students in U.S. colleges that would be all possible actions that are not in any formal syllabus or part of the required set of experiences in order to graduate.

When the literature in sociology started in the 1960's the term referred to the purely academic side of college. It was used to explain why first-generation students had lower grades and dropped out at a higher rate. In most of this paper we instead focus on actions students use during college to get hired into the career and position of their choosing<sup>3</sup>.

For a student to be affected by the hidden curriculum they also have to be either unaware or unfamiliar of it's existence. A necessary condition is to be the first-in-family to experience a certain organization.

With those two requirements in mind<sup>4</sup> we can now categorize some actions into being either part of the hidden curriculum or not.

A set of actions that are both formally required and experienced also by those whose parents stopped their schooling after high school, or an equivalent degree, are actions that are present in a class syllabus. GPA is a stated and formal requirement both of the university and by most jobs<sup>5</sup>. This means that actions that are clearly stated and that help increase grades will not be part of the hidden curriculum. One example of this

---

<sup>3</sup>This is a sound approach for two reasons: i) if modeled in it's entirety this step would come after a student solves a Roy model to determine career choice and is now optimizing how to reach that career, ii) the literature has shown that first-generation students are less likely to switch careers and majors (Aucejo et al. (2025)) therefore studying what they do after having picked a path is not only relevant but also theoretically sound.

<sup>4</sup>Unstated and never experienced by the student or their family.

<sup>5</sup>Most job applications for high skilled labor require submitting transcripts and GPA.



is doing class readings, doing an extra-hour of homework, or revising assignments when given the possibility.

On the other hand many high skilled jobs value signals of passion and preparedness that they do not formally state in the application process. For example cold-networking<sup>6</sup> or joining a student organization are both valued by hiring managers. However, they are never formally stated requirements from either the university or the job application.

## 2.2 Basic Framework

**Set Up** We start by building a basic model of skills investments that describes the relationship between students efforts and their outcomes. We will then introduce the friction of the hidden curriculum to explain the differences across students.

Students decide whether to invest a positive amount effort<sup>7</sup> into two separate inputs:  $h$  and  $v$ . The two inputs are respectively actions within the hidden and formal curriculum. The students decisions plus their baseline ability  $a$  that they have developed up until reaching college map into a quality index following

$$Q = f(h, v, a),$$

where  $f()$  is strictly increasing and concave in all three parameters.

Both the index  $Q$ , and all inputs are considered a scalar in this simple framework version. However, results generalize to cases where efforts  $h$  and  $v$ , and quality and ability measures,  $Q$  and  $a$ , are multidimensional<sup>8</sup>.

Each student wants to maximize expected future wages net of current costs of investments. The optimization problem therefore is the following

---

<sup>6</sup>Cold-networking is the action of reaching out and networking with someone you do not have a direct connection to. One very common case of cold-networking is reaching out to alumni from your institution and leveraging the indirect connection as a way to increase chances of positive engagement and interactions.

<sup>7</sup>This can be time, or can be seen as any form of exertion for example mental effort.

<sup>8</sup>In the case of multidimensional skills and multiple inputs we would follow the theoretical framework laid out by Cunha and Heckman (2007) and Deming (2017).

$$\max_{h,v} E[W(Q)] - C(h, v). \quad (1)$$

Wages  $W()$  are an increasing function of  $Q$ . The expectation operator in 1 captures uncertainty from future labor market shocks that are orthogonal to  $Q$ . It does not capture any uncertainty about how  $Q$  maps to wages.

**Mechanisms** There are two ways in which the hidden curriculum enters the students decision making problem as an information friction: actions  $h$  are not in the consideration set, and/or subjective beliefs of returns of actions  $h$  are lower for one group of students.

In our setting we have two groups of students.  $FG$  which are the uninformed, and  $CG$  which are the fully informed. We say that students of group  $FG$  suffer from the hidden curriculum if one of two things happens.

First in Equation 1 student  $FG$  will only maximize over  $v$  and will ignore  $h$ . In our model we interpret that as the student not knowing of the existence of  $h$ . Second we would see for same level of ability  $a$  and same mapping  $f()$  different beliefs on returns to investments in  $h$ . We model that as

$$\frac{\partial \hat{f}(h, v, a)_{FG}}{\partial h} < \frac{\partial \hat{f}(h, v, a)_{CG}}{\partial h}, \quad (2)$$

where  $\hat{f}()$  are the subjective beliefs of the student<sup>9</sup>.

The first-order conditions for the optimization problem lead to the following comparative statics.

*Proposition 1* (Hidden Curriculum Investment Gap): For students with identical ability  $a$  the equilibrium investments satisfy

$$h_{FG}^* \leq h_{CG}^*, \quad (3)$$

---

<sup>9</sup>In the parametric version of this model we assume CD function and the differences across students will be in the subjective beliefs surrounding the output elasticities.

if either  $h$  is not in the consideration set, or Equation 2 is true.

While we leave the proof in Appendix B the basic economic intuition is simple. If a student does not know of an input they mechanically will not invest in it. If they believe their returns are lower when investing in  $h$  compared to the beliefs of another student with all-else equal ability and costs, they will also invest less in it.

*Proposition 2*(Formal Investment Gap): The equilibrium investments in formal curriculum actions are a function of whether the two inputs are substitutes, complements, or independent. In the case of substitutes and independent inputs, then  $v_{FG} > v_{CG}$ . In the case of complements the direction is ambiguous and it would depend on whether the income effect or the substitution effect is stronger.

The economic intuition follows from two competing forces. First, when CG students invest more in hidden curriculum activities, they achieve higher quality  $Q$ , which reduces their marginal utility from additional investments due to concavity of  $W(Q)$  –this *income effect* pushes CG students to invest less in  $v$ . Second, if hidden and formal activities are complements in production, higher  $h$  makes  $v$  more productive – this *complementarity effect* pushes CG students to invest more in  $v$ <sup>10</sup>.

**Implications** For most settings everything else being equal both mechanisms imply higher  $v$  and lower  $h$ <sup>11</sup>. However, these two mechanisms imply very different learning dynamics and very different policy patterns.

As we will see in Section 5 students have a hard time investing time and effort in learning something they do not know about. This would imply that if most of hidden curriculum is coming through the consideration set mechanism, we would see slower learning across a students tenure in college. While on the other hand if the issue was mainly subjective beliefs, the student could, if incentives were aligned, invest effort into learning more about a certain action.

---

<sup>10</sup>The empirically relevant case appears to be substitutes or independence. When networking, office hours, and career preparation (hidden curriculum) can partially substitute for exceptional coursework performance, or when these activities independently contribute to quality, FG students compensate for their lack of hidden curriculum knowledge by working harder on visible, formal requirements. This generates a testable prediction: we should observe FG students spending more time on traditional coursework and assignments relative to their CG peers, consistent with qualitative evidence of first-generation students working hard but achieving similar or worse outcomes. Under complementarity, the prediction is ambiguous.

<sup>11</sup>If we added value of leisure or particular cost functions we could see not only a substitution pattern across  $h$  and  $v$  but also something akin to an income effect: therefore both inputs could be lower even for some ability  $a$ .

In terms of policy intervention to close the information gap the solution to a consideration set issue would not simply be to offer student resources. If the student does not know they lack knowledge of inputs, then they are less likely to invest time and effort into figuring out how and when to interact with on campus services such as first-generation offices and career services.

Therefore, understanding which mechanism is mainly to blame for the information friction is essential in order to both understand learning patterns in the data, but also to give actionable and scalable solutions to institutions.

### 3 Evidence from Observational Data

In this section, we will use a large survey linked with administrative data to show motivating facts on first-generation students and how their behaviors cannot be fully characterized by preferences and constraints. In particular, we will first show differences in outcomes similar to those in the literature. We will then show differences in effort allocated across the formal versus hidden curriculum. Finally, we will show when these differences arise and that there is little converge across cohorts over time.

#### 3.1 SERU Data

This section uses data from the “Student Experience in the Research University” (SERU) Consortium, specifically focusing on the 2022 cross-sectional dataset. The SERU Consortium conducts large-scale surveys targeting undergraduate students enrolled in major research-intensive universities.

Our sample includes more than 100,000 students from twenty institutions, primarily characterized by their high research activity, substantial undergraduate enrollments, and distinguished academic programs. These institutions range geographically across various regions of the United States, enabling us to examine patterns broadly representative of large research universities. The universities in our sample are also on average highly ranked. According to nationally recognized rankings we have 11 schools that are ranked in the top 50 undergraduate programs in the nation, 17 in the top 100. We therefore view first-generation students in our sample as highly positively selected and possibly a lower bound to information asymmetry between them and their continuing-generation counterparts.

The dataset includes extensive student-level survey responses, complemented by admin-

istrative data on student academic and demographic characteristics, such as cumulative GPA, high school GPA, tuition payments, credits taken, residency status, family social class, race, and gender.

Of particular interest for our analysis are variables directly linked to the hidden and formal curriculum. The dataset includes many variables related to actions within purely academic settings, such as various measures of faculty interactions (e.g. how many faculty they feel comfortable enough asking a reference letter), participation in study groups, and seeking help from faculty or peers. Additionally, the dataset includes actions that are extra-academic, such as peer help with networking, participation in internships, joining professional clubs, and taking on leadership roles. Given the extensive nature of the dataset in terms of inputs, in the next sections we will focus on a subset of actions – namely those survey items which are included at most schools.

As shown in Table 1 provides summary statistics comparing first-generation (FG) and continuing-generation (CG) students, highlighting key differences across observable characteristics. Notably, first-generation students tend to be older and are substantially more likely to come from lower-income and working-class backgrounds compared to their continuing-generation peers. Additionally, first-generation students are more frequently Hispanic or Latino, 47% to 14%, whereas continuing-generation students are disproportionately White, 46% compared to the 15% in first-generation, and first-generation students are also more likely to be from in-state 86% to 77%.

Additionally, the dataset contains optional survey modules that further enrich our understanding of student engagement, experiences, and behavior. While our main empirical analysis focuses specifically on measures directly related to the hidden curriculum and academic investments, additional variables capturing broader dimensions of student well-being, overall satisfaction, and college experience are available and discussed in supplementary analyses provided in Appendix A.

While, the availability of detailed demographic and academic background information allows us to condition our analysis on critical observables, enhancing our ability to rule out alternative explanations and provide robust insights into the relationship between first-generation status and student investments in formal and informal educational activities, we still view this section as motivational and not as proof of any causal link between first-generation status, information, and differences in effort choices.

## 3.2 Results

**Differences in Outcomes** We now move onto characterizing differences in outcomes between first-generation and continuing-generation students. We first highlight that in our sample first-generation students perform worse, and then show motivating evidence that constraints are not the sole cause of these differences.

We start by looking at college GPA. Continuing generation students have on average a 6.7% higher GPA than their first-generation peers. To test the robustness of these differences we regress college GPA on first-generation status. In the first of our two preferred specifications we only include controls on gender and ethnicity. In the second specification we add controls for social class, in-state residency, and whether they are a U.S. national or not. Both specifications include university fixed effects.

In Table 2 we report the results of an OLS on undergraduate GPA. We see a coefficient of  $-.13$  ( $p=0.000$ ) and  $-.075$  ( $p=0.000$ ) on the first-generation indicator, which corresponds respectively to one-fifth and one-seventh of a standard deviation.

While our controls may reduce a lot of worries about constraints that are common in the literature – e.g. financial constraints or language barrier driving the differences in outcome – we may still think that the main constraint is baseline ability. To control for such difference we use high school GPA<sup>12</sup> as a proxy for college preparedness.

First, as a sanity check we report in Table A.2 a regression of current GPA on various characteristics. The coefficient on high school GPA is not only highly significant but also has the largest value. We then include high school GPA as a control in our main specifications. We report small changes in the coefficients, indicating that baseline high school preparation does not explain much of the first-generation gap.

We then report regression results from the same specification as in Table 2 but for high school GPA as the outcome variable. We notice that the coefficient on being first-generation has a very small effect on high school GPA when compared to the coefficient on college GPA (Appendix Table A.3).

Now we pick the most conservative approach we can to control for constraints. We reduce our sample to: first years, from the U.S., who are upper middle or higher income, non-transfer students. We chose this sample for various reasons. U.S. national and higher income not only controls for financial resources and language barriers, but also allows us

---

<sup>12</sup>For high school GPA we have to drop observations from students that do not have traditional GPAs: e.g. out of 100, 60, and so on, since we cannot weigh them properly.

to proxy for high school quality. We pick first-years because our results may be driven by some unobserved heterogeneity moving students to learn different things during their time in college.

In Table 3 we see that first-generation status has no significance on high school GPA. However, it has a large negative association with current end of first year undergraduate GPA. A single year in college moves otherwise similar students to perform worse. We interpret the above results as an indicator that something other than ability, different demographic characteristics, or financial constraints are driving the divergence in GPA once they join college.

We now turn to credit and non-credit bearing internships<sup>13</sup>. In some sense internships could be considered part of the set of inputs a student uses to secure a full time job after graduating. However, given how similar the recruiting processes are and the increase in companies considering internships as a necessary step in order to secure their full time role, we include it as an outcome for this analysis.

We find that only 17% of first-generation students in our data report having ever secured an internship, compared to circa 20% of their continuing-generation peers. The difference is even more stark for non-credit bearing internships where 17% of first-generation students having completed one vs 26% of continuing-generation. As was the case for GPA, these differences are robust to all our controls and specifications (Table 4, and Table 5).

Especially important are the controls on social class for the results on non-credit internships. One could worry that not being for credit correlates with being unpaid. If this were the case then it would make sense for our first-generation students to be less likely to complete one, since on average they are more likely to be part of lower and working class families.

To eliminate this concern, we look at the difference across first-generation and continuing-generation uniquely within middle, upper-middle class and wealthy students. We report the results in the Appendix (Table A.4) and notice how there is a 9 percentage point difference ( $p=0.000$ ) in the two groups, which is circa one-fifth of a standard deviation.

**Differences in Actions** We take the above results as a potential indicator that first-generation students differ from their continuing-education peers in ways other than constraints. Our theory would predict that the difference comes from knowledge on what

---

<sup>13</sup>A credit bearing internship is an internship that counts towards a college credit.

actions/inputs to take during college in order to maximize success in the labor market. If this mechanism drives the observed patterns, first-generation students would report lower effort in hidden curriculum activities and higher effort in formal curriculum activities.

We pick a mix of academic and non-academic actions in both sets to show that this phenomena is true across domains. For the hidden curriculum actions we start with: i) number of faculty they talk to for strategic reasons<sup>14</sup>, and ii) whether they participate in student organizations. For the formal curriculum actions we look at: i) how many times they have presented in class, ii) how many of the class readings they have completed, iii) if they have enrolled in a methods class, iv) whether they have done an independent study course.

First-generation students have 14% less faculty to ask a reference from ( $P=0.000$ ), are 26% less likely to join a student organization ( $p=0.000$ ). From Table 6 we see that these differences are stable to adding controls and university fixed effect<sup>15</sup>. On the other hand we see that first-generation students do more class readings, more presentation, more independent research programs and are more likely to be enrolled in a methods class. Again these differences are robust to controls and university fixed effect (Table 7 and Table 8).

However we notice that these differences could derive from unobserved preferences and not from strategic behavior. For example continuing-generation students may like joining clubs more than first-generation students. If that is the case we should see some difference across the two groups even when the returns are flat. The reasoning is that if preferences are driving participation in club activities or other extra-curricular, we should already see a difference in engagement across the two groups even when it is not so efficient to do so.

We therefore pick two more non-required activities from the hidden curriculum. This time we pick activities that are major-specific, in the sense that only students from certain majors would get returns from allocating effort into these activities. We pick on campus entrepreneurial opportunities and leadership programs. These two actions should not matter for all students across all majors: for example a chemistry students interested in a career in academia may not find any returns from allocating effort into either.

In columns 2 and 5 of Table 9 we see how once we add all controls the difference between the two groups is not statistically significant. However, if we add in major fixed effect we

---

<sup>14</sup>With strategic reasons we mean any reason that is not directly to clarify a question from class, for example to receive a reference.

<sup>15</sup>These results are also robust to alternative specifications (Logit, Probit, and OLS).



find again a difference between the two groups. While this cannot rule out any form of preference for certain activities (and with the caveat that these are equilibrium outcomes), this helps us understand that the differences only arise when they matter and are not true across the board.

**Learning the Hidden Curriculum** We now focus our attention on how students differ across years. While our data does not allow us to have a panel-like structure we intend to give a snapshot of how students in different years behave. The reasoning is that even if students differ because of cohort characteristics they should still show some form of converging to similar means. Four years of college should be enough time to see behaviors of first-generation students converge to the means of their continuing-generation peers, even without fully ever reaching it due to unobservable characteristics.

In Figure 1 we report differences in OLS coefficients with university fixed effects across all hidden curriculum actions between first-generation and continuing-generation students. We look at first through fourth year for all non-transferred students.<sup>16</sup> We see very little convergence in behavior. For number of faculty we see a monotonic divergence and for entrepreneurial activities we see a divergence after year two. We see some small convergence in joining student organizations, however the gap never comes close to closing to zero.<sup>17</sup>

**Putting Everything Together** Our results from this section show that first-generation students have worse outcomes, they perform more actions that are formally defined in their curriculum, and perform less actions that are part of the hidden curriculum. We also showed that traditional controls do not explain the differences across the two groups. In Appendix Table A.5, we also show hidden curriculum actions accounts for the largest part of the  $R^2$  on whether first-generation students get an internship, even when compared to traditional controls. In a Oaxaca-Blinder decomposition exercise on getting a for credit internship, the number of faculty and whether the student is part of a students organization are the top two observables in explaining the difference between the two groups. As reported in Appendix Figure A.1 they explain 48% of the gap.

Finally as another robustness check we report Appendix Figure A.2 that there is a

---

<sup>16</sup>We cannot determine tenure of transfer students within their current institution.

<sup>17</sup>The locus of the Y axis may throw off the reader in believing first-generation students are converging to parity in student organizations. The difference only goes from circa -.17 to -.12, which is still very far from the 0 line.

positive correlation between all but one action<sup>18</sup> and getting an internship.

## 4 Information Provision Experiment

The administrative data analysis in Section 3 allows us to document substantial differences in Hidden Curriculum (HC) knowledge between FG and NFG students and rule out several potential explanations, however the data can only bring us so far. To pinpoint the causal link between information on actions and first-generation students choices in investments, we move to a framed field experiment (Harrison and List (2004)). This allows us to precisely control the assignment mechanism and better understand the mechanisms driving differences in HC knowledge.

### 4.1 Experimental Design

**Sample** To conduct the framed field experiment, we recruited students through two methods to reach a diverse, representative group of participants. The first method was by distributing the survey across campus with the help of a large group of majors offices and students organizations. We also publicized the survey through mailing lists and on-campus flyers. The second was by distributing the survey through the Xlab Berkeley (Experimental Social Science Laboratory). We report pooled results since student demographics do not look significantly different by distribution method<sup>19</sup>.

We screen students to make sure they are primarily interested in careers in business, finance, or technology.<sup>20</sup> The final sample consisted of 66% female students, 25% first-generation students, and 20% transfer. The two largest ethnic groups were Asian with 70% and White with 27%. The former group is over-represented in our experiment compared to the overall underlying population. However, we do not know the distribution of ethnicity for those interested in the careers we select on.

---

<sup>18</sup>Class readings have a negative relationship with getting an internship. We do not interpret this as a sign that readings are bad for getting a job, just that in equilibrium students that perform more readings are on average getting less internships. This could be because they already selected into a more academic track and therefore do not need an internship.

<sup>19</sup>Specifications that include distribution method fixed effects do report the same results.

<sup>20</sup>We actively excluded students interested in an academic career because our outcome variable of interest for this specific design would not have been natural for such students. However, future research should also include students interested in such careers since similar achievements gaps between first-generation and continuing-generation that have motivated this paper have been documented in academia and particularly in economics. Stansbury and Rodriguez (n.d.)

**Design** Upon recruitment, students completed an initial survey collecting baseline information and beliefs about effective strategies for securing employment post-graduation. This baseline elicitation served as a subtle mechanism to introduce the subsequent informational treatments without generating undue experimenter demand.

Participants were randomly assigned –using complete randomization – to one of three groups: a control group, a light touch information treatment, and a detailed and more heavy handed information treatment. In Table 10 we note no differences in observables across experimental groups.

The control group did not receive any informational intervention and was simply asked to report their beliefs about effective strategies to secure a job. The first treatment group (T1) received information emphasizing networking as a strategy to secure employment, designed to increase the salience of networking in their consideration set. The second treatment group (T2) received more explicit and detailed information about the returns and practical importance of networking, including statements such as: “Engaging with alumni can give you a competitive edge and help you get noticed in a crowded job market,” and “The best candidates often come from referrals”. We also highlight how there is evidence that it is a requirement for some elite firms to perform cold networking.<sup>21</sup>

**Outcomes** Our primary outcome measure is participants’ willingness-to-pay (WTP) for a networking-related service. Specifically, we asked students if they were willing to pay \$5 for a service that would provide them with five alumni from their institution working in their desired career fields. This measure serves as a behavioral proxy for students’ valuation of cold-networking as a career success strategy. We use cold-networking for this experiment for two reasons. First, for the careers of the students in the sample it is a necessary condition in finding a job, therefore it has the potential of being undervalued by those that are not familiar with the inner workings of such high-skilled and high-paying jobs. This highlights the subjective beliefs channel<sup>22</sup>. Second, disadvantaged students may only think of networking as an activity performed by those with existing connections: i.e. a family member or a friend will introduce you to someone. This may induce first-generation students to not know if the existence of cold-networking as an acceptable strategy, which highlights our second mechanism.

---

<sup>21</sup>This information comes from qualitative interviews with hiring managers and from the literature (e.g. Rivera (2016)).

<sup>22</sup>The fact that this is a requirement in these high skill high paying jobs also eliminates ambiguity around optimal approach. The action itself is optimal and the only consideration the student has to make is whether their outside option is better or not.

We designed the two treatment groups with the idea of teasing out the two mechanisms we introduced in the model. The light touch information intervention (T1) is designed to mainly load on the consideration set mechanism. While, the heavy handed information intervention (T2) loads heavily on both mechanisms. However, we cannot rule out any of the two mechanisms are at play in both T1 and T2 because by implying that a strategy is available we may also be updating beliefs on returns, and by heavily updating beliefs on returns through T2 we are increasing likelihood of the action being in the consideration set.

**Bounded Treatment Effects** We believe our results to be a lower bound of the issue of hidden curriculum for two distinct reasons.

First, even though we do not explicitly use the word networking in the control group, offering access to 5 contacts could induce students to think of networking, therefore including it in their consideration set. Therefore, the treatment effect we find is both a lower bound to the total effect of not having an action in your consideration set, but it is also mainly identifying the difference between knowing and not-knowing that cold-networking is an acceptable strategy.

Second, Berkeley students are highly positively selected. It therefore stands to reason that their first-generation students are some of the most informed when compared to the average or marginal first-generation student in U.S. colleges.

## 4.2 Experimental Results

**Baseline Differences** We begin by documenting substantial baseline differences in networking valuation across student types.<sup>23</sup> We notice differences in willingness to pay for the networking service by student first-generation background in the control group. First-generation students show a willingness to pay rate of circa 28% compared to continuing-generation students that exhibit a 40% willingness to pay rate.

These differences alone could imply credit or preference constraints from the first-generation students. We now look at whether treatment had any effect in order to identify the causal impact of information.

---

<sup>23</sup>Throughout this section we present OLS results for ease of discussion. However, in the appendix we present results for Probit and Logit regressions and we note no difference in either the magnitude of results or statistical significance.

**The Role of Information** We first look at the effects of the light touch information treatment (T1). As a reminder the light touch treatment (T1) emphasized networking as an available strategy, hence mainly loading on the consideration set mechanism.

We report that for first-generation students, T1 increased willingness to pay by 33% or 10 percentage points. Table 11 presents the primary experimental results. We regress whether the student is willing to pay the \$5 on a first-generation dummy, a treatment-dummy, and the interaction term between the two. In column 2 we add controls for race, GPA, gender, Internship and career of interest.

As expected from our discussion of baseline means, the coefficient on first-generation students is negative, economically large and statistically significant. The interaction term between the treatment and first-generation status is positive, statistically significant, and large enough to decrease baseline differences, such that the two groups are now not statistically different from each other.

This result indicates that in our sample most of the differences in behavior at baseline can be attributed to a lack of hidden curriculum activity within the consideration set.

We now look at the more heavy handed treatment, which provided explicit information about networking returns and necessity for competitive positions, generated a different pattern of responses. Table 12 presents the results from the same specification used for T1. Here we still notice a large positive impact from the T2 on first-generation students, (coefficient=.13). However, this is not statistically significant ( $p=0.2$ ).

While small sample size is contributing to the lack of statistical significant we also speculate that such a heavy handed treatment is backfiring on those individuals that already had the action in their consideration set but had very low subjective beliefs on the returns. The reasoning being that our treatment is not only trying to change beliefs about individual specific returns from networking but the necessity of cold-networking to get access to such high paying jobs. If a student has not done the action till now they may perceive that this information is too late for them, and they may be even more discouraged than before to take up the action.

**Other Demographics and Transfer Students** To test the robustness of our findings we compare our results from our first-generation students to those of other groups. In particular we look at two types of heterogeneities: i) groups that should not have this specific form of informational friction but that are usually disadvantaged in higher education, and ii) a group that instead could potentially suffer from the hidden curriculum.

For the “traditional” groups we look at gender<sup>24</sup> We do not find any significant effect in gender for either of the two treatments.

For the “other” group that may be affected by the hidden curriculum we look at transfer students from two-year community colleges. These individuals may have acquired the insider knowledge of the hidden curriculum specific to their old institutions. However, moving to a highly ranked four year college implies working with a completely new set of rules in the hidden curriculum.

For transfer students we report a baseline WTP of 23% which is 43% lower than non-transfer students. We find small treatment effects from T1 on transfer students. As reported in Appendix Table A.6 these differences are not statistically significant. On the other hand in Appendix Table A.7 we see that T2 has large and statistically significant effects on the WTP<sup>25</sup>. This implies that for transfer students the information friction stems largely from lower subjective beliefs.

**Implications** The experimental results provide several key insights into the mechanisms underlying the differences in willingness to invest in actions in the hidden curriculum. First, substantial baseline differences in networking valuation confirm that information frictions contribute to hidden curriculum gaps. Second, the differential responses to light versus heavy touch treatments across different groups suggest that first-generation students are mainly unaware of the availability of certain actions.

By corroborating this result of first generation students on transfer students we also see that each disadvantaged group faces its unique challenges, with the latter group showing signs of lower subjective beliefs on the returns of such actions.

These results suggest that targeted interventions should account for students’ existing knowledge and institutional experience. When the friction is a lack of knowledge information campaigns will bring greater returns than providing access to on-campus and on-demand resources: e.g. career service office. This is because students that do not know of the strategy, and that do not know about their information friction, will not see as much value in directed information search efforts. On the other hand more detailed information about returns may be necessary for students with some prior exposure.

---

<sup>24</sup>For ethnic minorities we only have enough power to look at Asian-Americans. We do not see any difference there as well.

<sup>25</sup>As for first-generation students, the treatment effects cancel out the baseline differences.

## 5 Search and Learning – Evidence from AI Tools

In the previous sections, the randomized information experiment established the existence of the hidden curriculum and that a lack of information and low subjective returns are the key mechanisms driving the gap in hidden curriculum knowledge. To better understand *why* this lack of information and perception of low returns persists, despite the seeming abundance of information and resources available to students at universities and online, we design a novel experiment leveraging AI tools to better understand how students search for and learn from information.

Studying information acquisition using interactive AI-powered chat-bots allows us to create a stigma-free, low-stakes environment where students can ask questions and search for information free from human judgment. This reduces potential barriers to asking for information or advice on sensitive topics and reduces concerns about signaling lower ability. Additionally, the use of a chat-bot allows us to track the path-dependent, sequential nature of information acquisition. Similar to real-life information acquisition, the dialogue with the LLM requires students to actively formulate their questions, process multiple complex and sometimes ambiguous responses, and make a conscious decision on the next question to ask and the next search step. Thus, the AI-tool closely mirrors real-world information acquisition.

Given the rise in AI-powered career information and education tools, we believe this is a realistic proxy for how students increasingly search for information (Freeman (2025); Attewell (2025)).

Finally, the heterogeneity in mechanisms driving the hidden curriculum also begs the need for a flexible policy tool in order for it to be scalable<sup>26</sup>. AI with its ability to adapt to the needs of the student and answer in a natural manner a variety of questions has the potential of being the scalable tool that universities and policy makers can use.

---

<sup>26</sup>The results on FG students would have pointed to an awareness campaign as being the best intervention. However, if then launched at scale for all disadvantaged group we would incur issues discussed in Al-Ubaydli et al. (2017). In particular we showed how each group may have unique needs, and this would create a drop in returns at scale.

## 5.1 Experimental Design

**Sample** We recruit 559 US college students using Prolific, an online survey platform.<sup>27</sup> Students first answer some basic demographics questions and a short survey. Students are then asked to “beta test” a new AI tool designed to help college students like them succeed in college and their careers. Our sample consists of 559 unique conversations, with  $N_{FG} = 217$   $N_{CG} = 342$  students distributed across the Control and Treatment arms<sup>28</sup>.

**Design** We use a  $2 \times 2$  design<sup>29</sup>. Students vary in first generation status (First-Generation (FG) versus Continuing-Generation (CG)). We then randomly assign students from both groups to the Treatment or Control AI tool.

The Treatment AI tool – or Active AI – is designed to answer questions that the students ask, while *actively* nudging students to consider Hidden Curriculum actions (e.g., professional networking, leveraging office hours, obtaining mentorship). The Control AI – or Passive AI – is designed to only answer with topics directly related to the question the student has asked. This allows for the possibility that it will nudge the student but to a much lesser degree than the Active AI.

The Active AI is designed with the goal to shock students consideration sets. We hope to mimic the effects in our information experiment with the added benefit that this is a targeted and flexible intervention: the information shock should be adaptive to the needs of the students. This is however not a mechanically true outcome. Students may not trust the AI or they may not know how to interpret the information provided. They also may be overly sure of their own priors and ignore any form of nudging.

The Passive AI is designed to understand what student’s search efforts would look like if they did not receive active guidance and had to decide their own path of discovery. By the nature of AI and by design we allow some variance in responses to see how amenable students are to nudges when their consideration set has not received a shock like for those students in the treatment group.

The two custom AI tools are created using an AI education platform currently used by

---

<sup>27</sup>We are also in the process of collecting a sample of UC Berkeley undergraduates through their Xlab network. In this section we only present results from the Prolific analysis as we are still in the process of collecting data from the Xlab sample. Results from the Xlab sample will be updated shortly.

<sup>28</sup>The difference in numbers for first-generation and continuing-generation students is not due to attrition but to a lower available sample of first-generation students.

<sup>29</sup>Even though there is only one randomization into either treatment or control, we call it  $2 \times 2$  for ease of exposition. Our analysis relies on comparing first generation students to their peers.



many educators including professors at UC Berkeley, the NYC Department of Education, and many others. We design the AI tools with low variability and largely restrict the information provided to students to the information we explicitly feed into the AI tools. This gives us a large amount of control over the information provided, allowing us to document differences in learning while holding the information itself largely fixed. We then use the transcripts from the chat history to quantify students’ consideration sets and learning efforts as they search for academic and career information.

**Outcomes** The key outcomes we are interested include:

*Initial Consideration Sets and AI Treatment Effect:* We analyze the share of student-initiated messages focused on Hidden Curriculum topics (or the User HC Rate). This measures the student’s starting consideration set and their willingness to seek information about hidden curriculum actions. We compare the incidence of HC topics between the Active AI Tool and the Passive AI Tool to measure the treatment effect of expanding the consideration set.

*Differences in Search Behavior:* We use LLM text analysis to identify identify patterns in how FG and CG students ask questions and search for information. Specifically, we are interested in the specificity of questions, the ability of the students to guide the conversation and extract relevant, actionable information from the LLM. We also document differences in language related to confidence and uncertainty.

*Nudges and Learning:* To better understand how learning happens, we proxy for a student’s willingness to engage in exploration on HC topics, we also track how often students “follow-up” on AI-initiated “nudges” on HC topics. We call this the HC Follow-Up Rate. A students has “followed-up” on a nudge if the student has asked a follow-up question or otherwise continued to discuss the HC topic following the AI Tool’s nudge. Combined with the baseline rate of HC topic discussion, this metric gives us a sense for how much the student is willing to engage with and learn about a HC topic.

## 5.2 Experimental Results

**Gap in Initial Consideration Sets** Figure 2 summarizes the experimental results. We see a significant baseline gap in the consideration sets between FG and NFG students. At baseline, FG students have a total of circa 11% of messages focusing on HC topics versus 16% for CG students. In an interacted regression model<sup>30</sup> between FG status and

---

<sup>30</sup>With controls for gender and race.

treatment on percent of HC messages, we find that FG indicator alone is negative and statistically significant ( $-.05, p = .04$ ). This 31% gap is economically significant and mirrors the differences we see in Section 3 and Section 4.

The AI Treatment, which actively injected HC topics into the conversation, substantially increased the incidence of HC topics for both FG and CG students ( $\beta_{Treat} = 0.13, p = .000$ ). The average User HC Rate for FG students rose to 0.25. The average User HC Rate for CG students rose to 0.29. The interaction term ( $FG \times Treat$ ) was small and not statistically significant ( $.02, p = .5$ ).

This result shows that in pure effort terms when the barrier of discovery (i.e., expanding the consideration set) is overcome, FG students engage in discussions about hidden curriculum topics at a higher rate. However a simple mean on what topics are discussed can hide mechanisms about learning and belief updating. In the next paragraphs we will look into both.

**Differences in Search Process and Uncertainty** We use text analysis from an LLM to document differences in how FG and CG students ask questions and communicate with the LLM. We find that questions from FG students tend to focus on career paths from specific majors, concerns about their competitiveness, specifically about GPA requirements, and practical, but general, advice such as questions about the importance of internships.

On the other hand, CG students tend to focus on questions specific to the job application process, such as networking and interview prep. CG students are also more likely to ask for *specific* and *personalized* advice – that is, they explain their exact situation to the AI tool before asking for advice to ensure the advice is tailored to them. Finally, CG students are also more likely to ask for specific resource, leading to concrete next steps.

We also find that FG students often ask more vague, less directed questions and frame their questions in a way that suggest less confidence and greater uncertainty. For example, when presented with a HC topic FG students were more likely to ask questions such as “Is networking really necessary for my major?” or express apprehension (“I have [networked] but I’m a little scared.”) FG students often sounded unsure both when initiating questions and when responding to HC nudges, expressing higher levels of anxiety overall.

The results suggest that for students who are not part of the established network (FG students), the implicit cost of discovering and formulating an effective query about a hidden, or non-obvious topic – like the specific value of building a relationship with a non-teaching professor or how to secure an off-cycle internship – is substantially higher.

They often do not know the “right” questions to ask, therefore their information search is often less effective. This systematic lack of initial exploration underscores the finding that the action space for FG students is initially constrained not by ability or motivation, but by the absence of information on the cultural scaffolding that defines which topics are valuable enough to merit search.

Additionally we want to highlight the importance of AI as a tool for qualitative analysis. By going beyond the quantitative measures of whether or not they asked about a HC actions, from the qualitative analysis we see that FG students may be more reticent in updating their beliefs than the raw numbers would lead us to believe<sup>31</sup>.

**Endogenous Updating: Limited Learning in the Face of Information** While FG students were successfully brought to the frontier of HC topics by the AI, a significant difference emerged in their subsequent learning effort (persistence) when presented with a new HC concept via an AI nudge. In Figure 3 we analyze the rate at which students followed up on the AI Tool’s nudge toward a HC topic. FG students in the control group had a 25% HC Nudge Follow-Up Rate while CG students in the Control group had a significantly higher HC Nudge Follow-up Rate of 48% (a gap of  $-23$  percentage points).

However, the Treatment condition dramatically changed this learning effort gap. The HC Follow-Up Rate for FG increased to 75% while the CG Follow-Up rate increased to 54%. This drastic change in behavior suggests that CG students were closer to their frontier of interest for hidden curriculum topics. While on the other hand FG students were very far away. Once they receive this shock in consideration set they not only increase their effort because they are now interested in the topic but they also have to catch up on the details of how to mechanically implement the new advice.

This increase also suggests that the comprehensive information provided by the LLM in the Treatment arm acted as a powerful commitment device, boosting the perceived returns to learning and encouraging deeper engagement. The LLM essentially acted as an information broker, converting a vague, high-ambiguity signal into a concrete, low-risk action plan, which drastically lowered the threshold for effort investment for FG students. The magnitude of this effect –nearly tripling the follow– up rate for FG students–suggests that once the initial uncertainty is dispelled, their appetite for learning and investing effort is extremely high<sup>32</sup>.

---

<sup>31</sup>This could be due to lack of trusting AI, tight priors on their own beliefs, or simply less experienced at using AI.

<sup>32</sup>We do however want to bring back the attention to our qualitative results that suggest that FG

The large differences in HC Follow-Up rates between the Treatment and Control condition shows that FG students are far less likely to commit to learning about HC topics, even when minimally exposed. This aligns with our prior experiment’s finding of low subjective beliefs on returns – if the return is uncertain, the optimal learning effort is low. The 23 percentage point difference uncovers a clear difference in the perceived value of investment in effort; FG students rationally choose to allocate scarce cognitive resources elsewhere, given lower levels of starting confidence and the risk and ambiguity of returns to HC topics.

Overall, the results suggest that the frictions in information acquisition for FG students is rooted in both differences in their initial consideration sets size, how they search for new information, and how much they update in response to the new information, driven in part by their initial uncertainty and lack of confidence. These findings help us better understand the mechanisms of our first experiment by showing how differences in the hidden curriculum gap may persist over time.

## 6 Conclusion

This paper reveals how the hidden curriculum perpetuates inequality even when formal barriers to opportunity have been removed. Our evidence shows that unwritten rules and informal strategies create systematic disadvantages for first-generation students that compound over time, contributing to persistent gaps in economic and social outcomes.

Our findings challenge conventional explanations for inequality. The differences we document between first-generation and continuing-generation students cannot be explained by ability, financial constraints, or preferences. Instead, they stem from differential access to tacit knowledge about how institutions actually work. This knowledge that advantaged students inherit through family networks while disadvantaged students must discover through costly trial and error, if at all.

Three aspects of our results are particularly striking for understanding inequality. First, the hidden curriculum creates a perverse dynamic where those who work hardest may achieve least. First-generation students compensate for their lack of informal strategies by overinvesting in formal requirements, doing more readings, presentations, and coursework, and yet still achieve lower GPAs and fewer internships. This mismatch between effort and

---

students do not leave the experiment fully convinced and knowledgeable on what to do. This is a first step on how to solve the problem: not a one stop shop to fully closing the HC gap.

reward violates basic notions of meritocracy and fairness.

Second, our experiments shows that these inequalities are surprisingly easy to perpetuate yet potentially simple to address. A brief information intervention essentially eliminated the first-generation gap in networking beliefs, suggesting that vast inequalities rest on thin informational foundations. The problem is not that disadvantaged students lack ability or motivation, but that success requires knowledge of rules that are never explicitly stated.

Third, the AI experiment reveals a troubling feedback loop. Students who most need information about hidden curriculum strategies are least able to find and process it effectively. They ask different questions, express more uncertainty, and are less likely to follow up on new information. This creates an inequality trap where initial disadvantages in cultural capital translate into persistent differences in information acquisition, making catch-up increasingly difficult over time.

The implications extend far beyond individual students. When capable first-generation students systematically miss opportunities due to information they were never given, society loses potential innovations, diverse perspectives in leadership, and the social mobility that justifies unequal rewards.

The path forward requires recognizing that equal access is not equal opportunity when success depends on unwritten rules. Institutions serious about equity must make the hidden curriculum visible. However, our results on heterogeneity of mechanisms imply that each disadvantaged group would need it's own tailored intervention. For example groups affected by consideration set issues will not be moved by opportunities to engage with on-campus offices and resources, since they are not aware of their own lack of knowledge.

Future work should examine how hidden curricula operate in other domains: workplace advancement, entrepreneurship, political participation. Understanding these mechanisms across contexts could inform broader efforts to create genuinely equal opportunity. The hidden curriculum represents a fundamental challenge to meritocratic ideals. By documenting its existence and effects, we hope to spark conversation about the gap between formal equality and substantive opportunity. Only by making the invisible visible can we begin to address the subtle mechanisms that perpetuate inequality across generations.

## References

- Al-Ubaydli, Omar, John A List, and Dana L Suskind**, “What can we learn from experiments? Understanding the threats to the scalability of experimental results,” *American Economic Review*, 2017, 107 (5), 282–286.
- Altmann, Steffen, Armin Falk, Simon Jäger, and Florian Zimmermann**, “Learning about job search: A field experiment with job seekers in Germany,” *Journal of Public Economics*, 2018, 164, 33–49.
- Arcidiacono, Peter, V Joseph Hotz, Arnaud Maurel, and Teresa Romano**, “Ex ante returns and occupational choice,” *Journal of Political Economy*, 2020, 128 (12), 4475–4522.
- Attewell, Sue**, “Student Perceptions of AI 2025,” May 2025. Report on how students use AI, including employability tasks (CVs, applications, interviews). Accessed October 8, 2025.
- Aucejo, Esteban M, Jacob French, Paola Ugalde Araya, and Basit Zafar**, “Understanding gaps in college outcomes by first-generation status,” Technical Report, National Bureau of Economic Research 2025.
- Ba, Cuimin, J Aislinn Bohren, and Alex Imas**, “Over-and underreaction to information,” *Available at SSRN 4274617*, 2024.
- Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakçı, and Rei Mariman**, “Generative AI can harm learning,” Technical Report, Wharton School Research Paper July 2024.
- Becker, Gary S. and Nigel Toms**, “An equilibrium theory of the distribution of income and intergenerational mobility,” *Journal of Political Economy*, 1979, 87 (6), 1153–1189.
- Belot, Michele, Philipp Kircher, and Paul Muller**, “Providing advice to jobseekers at low cost: An experimental study on online advice,” *The review of economic studies*, 2019, 86 (4), 1411–1447.
- Bourdieu, Pierre**, “The Forms of Capital,” in John G. Richardson, ed., *Handbook of Theory and Research for the Sociology of Education*, New York: Greenwood, 1986, pp. 241–258.

- Bowen, William G., Matthew M. Chingos, and Michael S. McPherson**, *Crossing the Finish Line: Completing College at America's Public Universities*, Princeton University Press, 2009.
- Bundorf, M. Kate, Maria Polyakova, and Ming Tai-Seale**, "How do humans interact with algorithms? Experimental evidence from health insurance," NBER Working Paper 25976, National Bureau of Economic Research June 2019.
- Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott**, "Misperceived social norms: Women working outside the home in Saudi Arabia," *American economic review*, 2020, 110 (10), 2997–3029.
- Calarco, Jessica McCrory**, "'I need help!'" social class and children's help-seeking in elementary school," *American Sociological Review*, 2011, 76 (6), 862–882.
- Calvo-Armengol, Antoni and Matthew O. Jackson**, "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 2004, 94 (3), 426–454.
- Caplin, Andrew, David Deming, Shangwen Li, Daniel Martin, Philip Marx, Ben Weidmann, and Kadachi Jiada Ye**, "The ABCs of Who Benefits from Working with AI: Ability, Beliefs, and Calibration," *NBER Working Paper*, October 2024, w33021.
- Chatterji, Aaron, Tom Cunningham, David Deming, Zoë Hitzig, Christopher Ong, Carl Shan, and Kevin Wadman**, "How people use ChatGPT," Technical Report, OpenAI September 2025.
- Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan**, "Mobility report cards: The role of colleges in intergenerational mobility," NBER Working Paper 23618, National Bureau of Economic Research 2017.
- , Nathaniel Hendren, Maggie R Jones, and Sonya R Porter**, "Race and economic opportunity in the United States: An intergenerational perspective," *The Quarterly Journal of Economics*, 2020, 135 (2), 711–783.
- Chopra, Felix and Ingar Haaland**, "Conducting qualitative interviews with AI," 2023.
- Croes, Esther A. J. and Marjolijn L. Antheunis**, "Digital confessions: The willingness to disclose intimate information to a chatbot and its impact on emotional well-being," *Interacting with Computers*, 2024, 36 (5), 279–297.

- Cullen, Zoë and Ricardo Perez-Truglia**, “The Old Boys’ Club: Schmoozing and the Gender Gap,” July 2022.
- Cunha, Flavio and James Heckman**, “The technology of skill formation,” *American economic review*, 2007, 97 (2), 31–47.
- Cunha, Flávio, Irma Elo, and Jennifer Culhane**, “Eliciting maternal expectations about the technology of cognitive skill formation,” Technical Report, National Bureau of Economic Research 2013.
- , —, and —, “Maternal subjective expectations about the technology of skill formation predict investments in children one year later,” *Journal of Econometrics*, 2022, 231 (1), 3–32.
- DellaVigna, Stefano**, “Psychology and Economics: Evidence from the Field,” *Journal of Economic Literature*, May 2009, 47 (2), 315–372.
- Deming, David J.**, “The Growing Importance of Social Skills in the Labor Market,” *The Quarterly Journal of Economics*, November 2017, 132 (4), 1593–1640.
- Dizon-Ross, Rebecca**, “Parents’ Beliefs about Their Children’s Academic Ability: Implications for Educational Investments,” *American Economic Review*, August 2019, 109 (8), 2728–2765.
- Dynarski, Susan M. and Judith Scott-Clayton**, “Financial aid policy: Lessons from research,” *The Future of Children*, 2013, 23 (1), 67–91.
- Ersoy, Fulya**, “Effects of perceived productivity on study effort: Evidence from a field experiment,” *Journal of Economic Behavior & Organization*, 2023, 207, 376–391.
- Freeman, Josh**, “Student Generative AI Survey 2025,” February 2025. Finds that 92% of students report using AI tools. Accessed October 8, 2025.
- Gabaix, Xavier**, “Behavioral inattention,” in “Handbook of behavioral economics: Applications and foundations 1,” Vol. 2, Elsevier, 2019, pp. 261–343.
- Graesser, Arthur C., Philip Chipman, Brian C. Haynes, and Andrew Olney**, “AutoTutor: A tutor with dialogue in natural language,” *Behavior Research Methods, Instruments, & Computers*, 2004, 36 (2), 180–192.



- Granovetter, Mark S**, “The strength of weak ties,” *American journal of sociology*, 1973, 78 (6), 1360–1380.
- Hafferty, Frederic W.**, “Beyond Curriculum Reform: Confronting Medicine’s Hidden Curriculum,” *Academic Medicine*, 1998, 73 (4), 403–407.
- Harrison, Glenn W and John A List**, “Field experiments,” *Journal of Economic literature*, 2004, 42 (4), 1009–1055.
- Hoxby, Caroline M and Christopher Avery**, “The missing” one-offs”: The hidden supply of high-achieving, low income students,” Technical Report, National Bureau of Economic Research 2012.
- Ioannides, Yannis M and Linda Datcher Loury**, “Job information networks, neighborhood effects, and inequality,” *Journal of economic literature*, 2004, 42 (4), 1056–1093.
- Jabarian, Brian and Luca Henkel**, “Voice AI in Firms: A Natural Field Experiment on Automated Job Interviews,” *Firms: A Natural Field Experiment on Automated Job Interviews (August 18, 2025)*, 2025.
- Jack, Anthony Abraham**, *The Privileged Poor: How Elite Colleges Are Failing Disadvantaged Students*, Cambridge, MA: Harvard University Press, 2019.
- Jackson, Philip W.**, *Life in Classrooms*, New York: Holt, Rinehart and Winston, 1968.
- Jensen, Robert**, “The (Perceived) Returns to Education and the Demand for Schooling \*,” *Quarterly Journal of Economics*, May 2010, 125 (2), 515–548.
- Kahn, Lisa B.**, “The long-term labor market consequences of graduating from college in a bad economy,” *Labour Economics*, April 2010, 17 (2), 303–316.
- Karna, Udit, John A List, Andrew Simon, and Haruka Uchida**, “Five Facts About the First-Generation Excellence Gap,” Technical Report, National Bureau of Economic Research 2025.
- Kestin, Greg, Kelly Miller, Anna Klates, Timothy Milbourne, and Gregorio Ponti**, “AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting,” *Scientific Reports*, 2025, 15 (1), 97652.

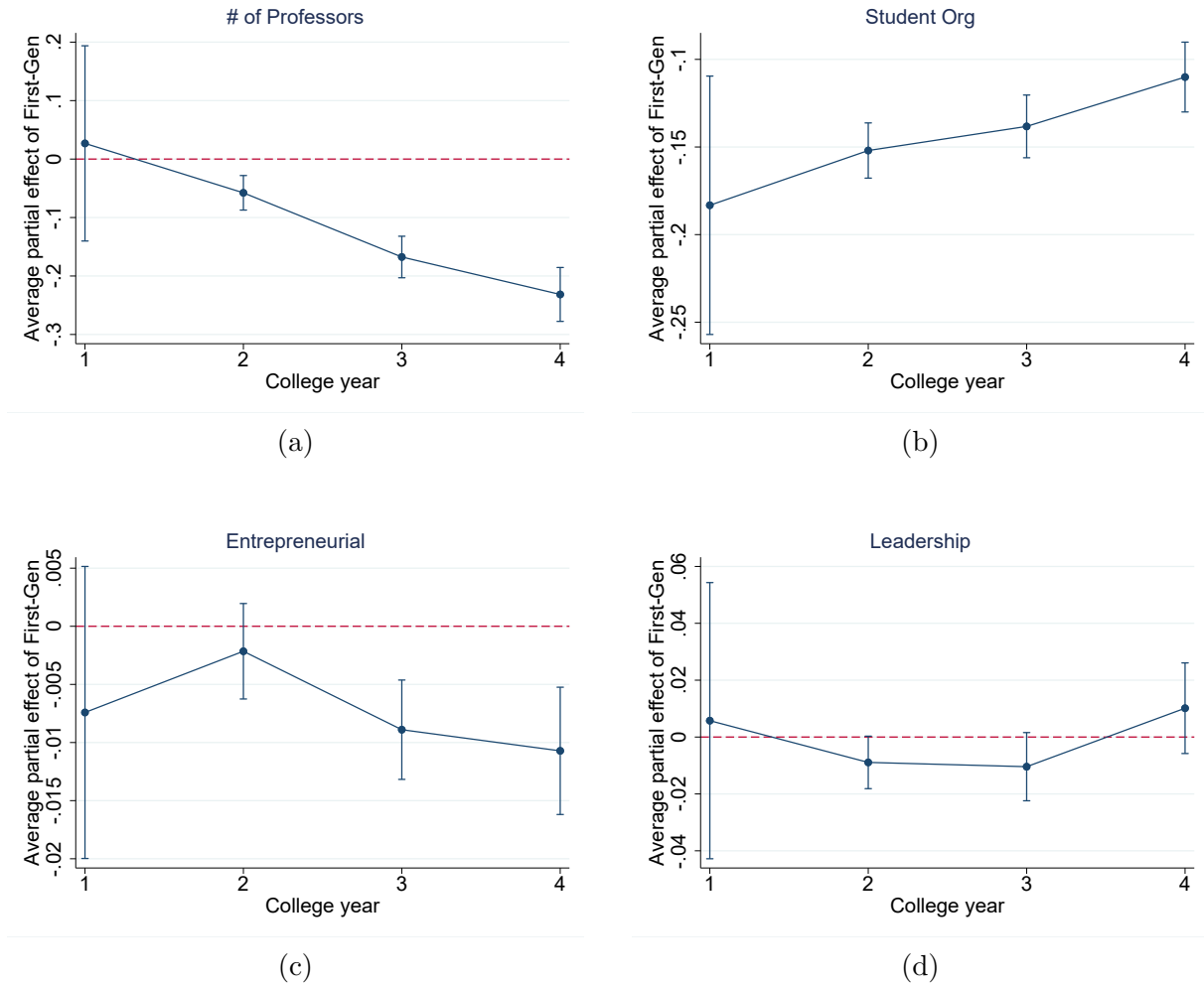
- Koedinger, Kenneth R. and Vincent Aleven**, “Exploring the assistance dilemma in experiments with cognitive tutors,” *Educational Psychology Review*, 2007, 19 (3), 239–264.
- , **Sidney K. D’Mello, Elizabeth A. McLaughlin, Zachary A. Pardos, and Carolyn P. Rosé**, “Data mining and education,” *Wiley Interdisciplinary Reviews: Cognitive Science*, 2015, 6 (4), 333–353.
- Lai, Vivian, Chacha Chen, Q Vera Liao, Alison Smith-Renner, and Chenhao Tan**, “Towards a science of human-ai decision making: a survey of empirical studies,” *arXiv preprint arXiv:2112.11471*, 2021.
- Lareau, Annette**, “Invisible inequality: Social class and childrearing in black families and white families,” *American Sociological Review*, 2002, 67 (5), 747–776.
- Lawrence, Carlton, Tsholofelo Mhlaba, Kearsley A. Stewart, Relebohile Moleletsane, Bernhard Gaede, and Mosa Moshabela**, “The Hidden Curricula of Medical Education: A Scoping Review,” *Academic medicine : journal of the Association of American Medical Colleges*, April 2018, 93 (4), 648–656.
- Lucas, Gale M., Jonathan Gratch, Andrew King, and Louis-Philippe Morency**, “It’s only a computer: Virtual humans increase willingness to disclose,” *Computers in Human Behavior*, 2014, 37, 94–100.
- Manzini, Paola and Marco Mariotti**, “Stochastic Choice and Consideration Sets,” *Econometrica*, 2014, 82 (3), 1153–1176. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10575>.
- Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y Ozbay**, “Revealed Attention,” *American Economic Review*, August 2012, 102 (5), 2183–2205.
- McCall, John Joseph**, “Economics of information and job search,” *The Quarterly Journal of Economics*, 1970, 84 (1), 113–126.
- Meyer, Johannes, Stephan Sting, Andreas Döring, Patrick Knorz, and Andreas Janson**, “Using llms to bring evidence-based feedback into the classroom,” *Computers & Education: Artificial Intelligence*, 2024, 5, 100264.
- Mo, Fengran, Kelong Mao, Ziliang Zhao, Hongjin Qian, Haonan Chen, Yiruo Cheng, Xiaoxi Li, Yutao Zhu, Zhicheng Dou, and Jian-Yun Nie**, “A survey of conversational search,” 2024. arXiv: 2410.15576 [cs.IR].

- Oyer, Paul**, “Initial Labor Market Conditions and Long-Term Outcomes for Economists,” *Journal of Economic Perspectives*, May 2006, *20* (3), 143–160.
- , “The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income,” *The Journal of Finance*, 2008, *63* (6), 2601–2628. Publisher: [American Finance Association, Wiley].
- Paruchuri, Akshay, Maryam Aziz, Rohit Vartak, Ayman Ali, Best Uchegara, Xin Liu, Ishan Chatterjee, and Monica Agrawal**, “What’s up, doc?": Analyzing how users seek health information in large-scale conversational AI datasets,” 2025. arXiv: 2506.21532 [cs.HC].
- Portelli, John P**, “Exposing the hidden curriculum,” *Journal of curriculum studies*, 1993, *25* (4), 343–358.
- Rivera, Lauren A**, “Pedigree: How elite students get elite jobs,” 2016.
- Roll, Ido, Vincent Aleven, Bruce M. McLaren, and Kenneth R. Koedinger**, “Can help seeking be tutored? Searching for the secret sauce of metacognitive tutoring,” in “Proceedings of AIED 2007” 2007, pp. 203–210.
- Rury, Derek and Scott E Carrell**, “Knowing what it takes: The effect of information about returns to studying on study effort and achievement,” *Economics of Education Review*, 2023, *94*, 102400.
- Solon, Gary**, “Intergenerational mobility in the labor market,” in “Handbook of labor economics,” Vol. 3, Elsevier, 1999, pp. 1761–1800.
- Stansbury, Anna and Kyra Rodriguez**, “The Class Gap in Career Progression: Evidence from US academia.”
- Stigler, George J**, “Information in the labor market,” *Journal of political economy*, 1962, *70* (5, Part 2), 94–105.
- Topa, Giorgio**, “Labor markets and referrals,” in “Handbook of social economics,” Vol. 1, Elsevier, 2011, pp. 1193–1221.
- Wiswall, M. and B. Zafar**, “Determinants of College Major Choice: Identification using an Information Experiment,” *The Review of Economic Studies*, April 2015, *82* (2), 791–824.

**Zafar, Basit**, “How do college students form expectations?,” *Journal of Labor Economics*, 2011, *29* (2), 301–348.

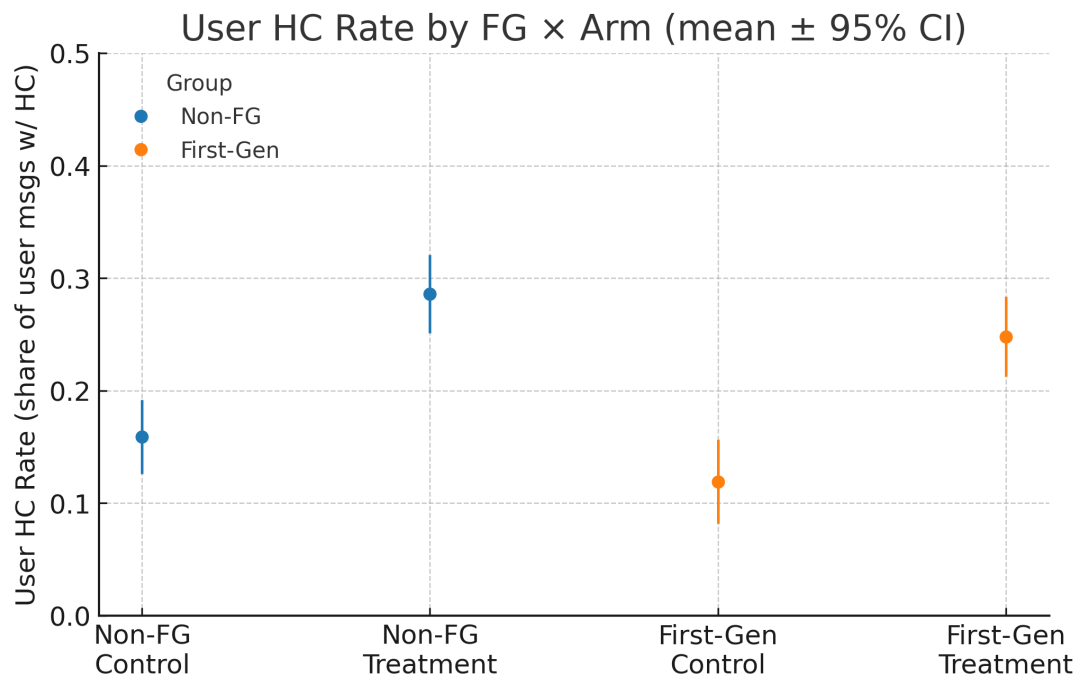
# Figures and Tables

Figure 1: Differences in Hidden Curriculum Levels



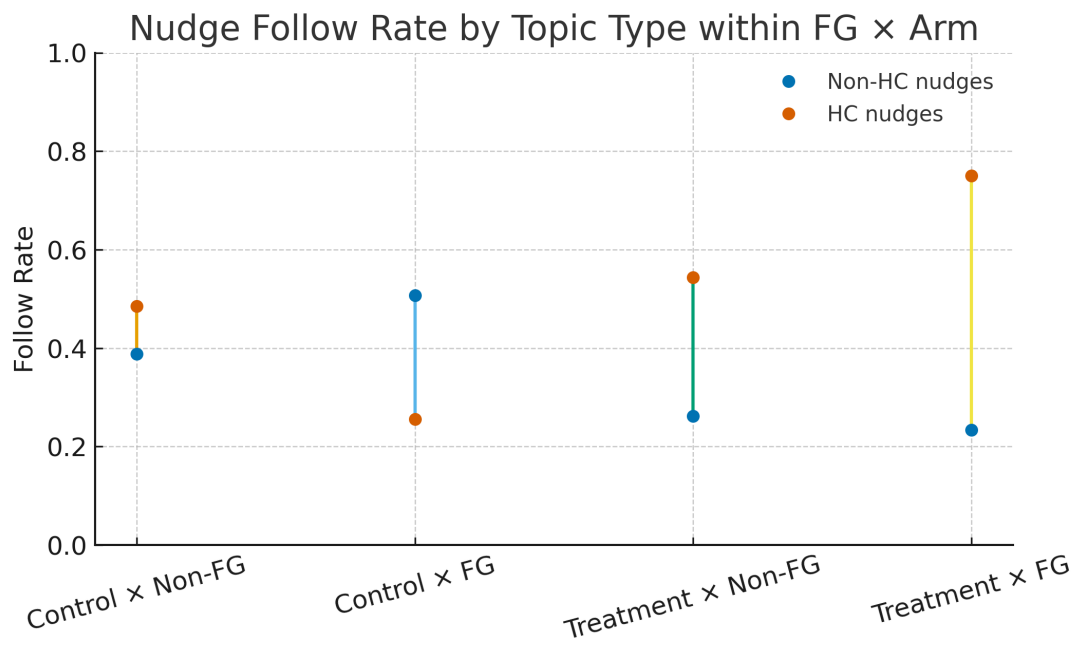
*Notes:* In all Panels we show the difference in OLS coefficients for HC activities between first-generation and continuing-generation students. The bars represent the 95% CI. In Panel 1a the outcome variable is the # of Professors the students network with. In Panel 1b the outcome variable is joining student organizations. In Panel 1c the outcome variable is official entrepreneurial activities. In Panel 1d the outcome variable is participating in leadership programs.

Figure 2: Differences in HC Topics



*Notes:* The figure shows means of user HC Rates across treatment group and FG status. HC rates are determined by classifying users messages as either a hidden curriculum topic or not and taking the ratio with respect to total number of messages sent by the user. AI messages are not included in this computation. Bars indicate 95% CI intervals.

Figure 3: Differences in Following Nudges



*Notes:* The figure shows means of user follow rates across HC status (whether message is HC topic or now), treatment group and FG status. Follow rates are determined by classifying the AIs message topic and then looking at whether the user followed up with a message on the same topic.

Table 1: Balance Table by First Generation Status

	First-Gen		Continuing-Gen		Diff
	Mean	(SD)	Mean	(SD)	
Age	21.36	( 4.71)	20.46	( 2.82)	0.90***
Female	0.67	( 0.47)	0.63	( 0.48)	0.04***
High School GPA	3.75	( 1.49)	3.94	( 3.19)	-0.18***
Low Income	0.40	( 0.49)	0.05	( 0.22)	0.35***
Working Class	0.35	( 0.48)	0.13	( 0.33)	0.23***
Middle Class	0.16	( 0.37)	0.37	( 0.48)	-0.21***
High Income	0.00	( 0.06)	0.03	( 0.16)	-0.02***
In-State	0.85	( 0.36)	0.74	( 0.44)	0.11***
White	0.17	( 0.38)	0.46	( 0.50)	-0.29***
Hispanic	0.47	( 0.50)	0.13	( 0.34)	0.34***
Asian	0.26	( 0.44)	0.26	( 0.44)	-0.00
Black/African American	0.04	( 0.20)	0.04	( 0.20)	-0.00
American Indian	0.01	( 0.09)	0.01	( 0.09)	0.00
Pacific Islander	0.00	( 0.05)	0.00	( 0.05)	-0.00
Race/Ethnicity Unknown	0.01	( 0.10)	0.03	( 0.17)	-0.02***
Observations	19,126		97,027		

*Notes:* Table 1 shows means, standard deviations (in parentheses), and differences across first-generation and continuing-generation students. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .



Table 2: First Generation on GPA (OLS)

	GPA at Current Institution		
	(1)	(2)	(3)
First-Generation	-0.198*** (0.005)	-0.128*** (0.005)	-0.075*** (0.006)
Observations	93,639	92,561	76,430
Controls 1	No	Yes	Yes
Controls 2	No	No	Yes
University FE	Yes	Yes	Yes

*Notes:* Table 2 shows results from OLS regression with university level fixed effects on GPA at their current institution. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Table 3: First Generation Effects in Early College

	Cumulative GPA	High School GPA
	(1)	(2)
First-Generation	-0.123** (0.051)	-0.016 (0.025)
Observations	3,354	4,498
Controls 1	Yes	Yes
Controls 2	Yes	Yes
University FE	Yes	Yes

*Notes:* Table 3 shows results from OLS regression with university level fixed effects. Sample is restricted to first-year, upper-middle or wealthy, U.S. national. In column (1) the outcome variable is current GPA and in column (2) it is high school GPA. Controls 1 are: gender and ethnicity. Controls 2 in this case is: being from in-state. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Table 4: First Generation and For Credit Internships

	For Credit Internships	
	(1)	(2)
First-Generation	-0.019*** (0.003)	-0.014*** (0.004)
Observations	110,250	95,613
Controls 1	Yes	Yes
Controls 2	No	Yes
University FE	Yes	Yes

*Notes:* Table 4 shows results from a Probit regression on whether student did a for credit internship. Entries are average marginal effects (AME) on the probability scale. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . For binary regressors, AME is the discrete change from 0 to 1.

Table 5: First Generation and No Credit Internships

	No Credit Internships	
	(1)	(2)
First-Generation	-0.065*** (0.004)	-0.047*** (0.004)
Observations	110,250	95,613
Controls 1	Yes	Yes
Controls 2	No	Yes
University FE	Yes	Yes

*Notes:* Table 5 shows results from a Probit regression on whether student did a not for credit internship. Entries are average marginal effects (AME) on the probability scale. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . For binary regressors, AME is the discrete change from 0 to 1.

Table 6: First Generation and Hidden Curriculum 1 (OLS)

	# of Faculty		Student Clubs	
	(1)	(2)	(3)	(4)
First Generation	-0.051*** (0.010)	-0.037*** (0.011)	-0.111*** (0.004)	-0.070*** (0.005)
Sample Size	104,754	96,853	111,104	96,393
Controls 1	Yes	Yes	Yes	Yes
Controls 2	No	Yes	No	Yes
University FE	Yes	Yes	Yes	Yes

*Notes:* Table 6 shows results from OLS regression with university level fixed effects on number of faculty a student talks to in order to get a reference and student club participation. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$

Table 7: First-Gen and Formal Actions 1 (OLS)

	Presentations		Readings	
	(1)	(2)	(3)	(4)
First Generation	0.060*** (0.012)	0.044*** (0.015)	0.060*** (0.010)	0.058*** (0.012)
Sample Size	114,129	96,714	112,666	96,894
Controls 1	Yes	Yes	Yes	Yes
Controls 2	No	Yes	No	Yes
University FE	Yes	Yes	Yes	Yes

*Notes:* Table 7 shows results from OLS regression on number of presentations and class readings. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Table 8: First-Gen and Formal Actions 2 (OLS)

	Methods Class		Ind. Research	
	(1)	(2)	(3)	(4)
First Generation	0.037*** (0.004)	0.021*** (0.005)	0.035*** (0.004)	0.030*** (0.004)
Sample Size	109,939	96,354	109,579	96,061
Controls 1	Yes	Yes	Yes	Yes
Controls 2	No	Yes	No	Yes
University FE	Yes	Yes	Yes	Yes

*Notes:* Table 8 shows results from OLS regression with university level fixed effects on whether students took a methods class or performed independent research. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$

Table 9: First Generation and Hidden Curriculum 1 (OLS)

	Entrepreneurial Activity			Leadership		
	(1)	(2)	(3)	(4)	(5)	(6)
First Generation	-0.003** (0.001)	-0.001 (0.001)	-0.003* (0.002)	-0.003 (0.003)	-0.002 (0.003)	-0.008* (0.004)
Sample Size	111,026	96,294	67,154	111,026	96,294	67,154
Controls 1	Yes	Yes	Yes	Yes	Yes	Yes
Controls 2	No	Yes	Yes	No	Yes	Yes
Major Controls	No	No	Yes	No	No	Yes
University FE	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* Table 9 shows results from an OLS regression with university level fixed effects on whether the student has performed Entrepreneurial activities or participated in leadership programs. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Table 10: Balance Table: Means and SDs by Group

	Control C		Treatment 1 T1		Treatment 2 T2		Difference $\Delta$	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	T1 - C	T2 - C
Female	0.66	0.47	0.64	0.48	0.67	0.47	-0.03	0.01
First-Gen	0.23	0.42	0.29	0.46	0.24	0.43	0.06	0.01
Transfer	0.18	0.38	0.18	0.39	0.23	0.42	0.01	0.05
Asian	0.69	0.46	0.67	0.47	0.67	0.47	-0.03	-0.03
Black	0.01	0.12	0.01	0.12	0.01	0.12	0.00	-0.00
White	0.25	0.44	0.28	0.45	0.28	0.45	0.02	0.03
Observations	213		212		220			

Notes: Table 10 shows means, standard deviations (in parenthesis), and differences across experimental groups. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 11: OLS Coefficient on Networking Email for First Generation Students

	(1) Networking Emails	(2) Networking Emails
Treatment 1	-0.075 (0.055)	-0.087 (0.055)
First-Generation	-0.123 (0.075)	-0.157** (0.077)
Treatment 1 x First-Gen	0.177* (0.105)	0.232** (0.105)
Controls	No	Yes
SE Type	Robust	Robust
Observations	425	425

Notes: Table 11 shows the outcome for whether or not a student is interested in receiving networking emails on first generation status. Entries are OLS coefficients. Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 12: OLS Estimation – Networking Email for First Generation Students

	(1) Networking Emails	(2) Networking Emails
Treatment 2	-0.016 (0.054)	-0.028 (0.054)
First-Generation	-0.123 (0.075)	-0.171** (0.078)
Treatment 2 x First-Gen	0.115 (0.108)	0.137 (0.110)
Controls	No	Yes
SE Type	Robust	Robust
Observations	433	433

*Notes:* Table 12 shows the outcome for whether or not a student is interested in receiving networking emails on first generation status. Entries are OLS coefficients. Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 13: OLS Estimation – Networking Email for Female Students

	(1) Networking Emails	(2) Networking Emails
Treatment 1	-0.038 (0.079)	-0.051 (0.080)
Female	-0.013 (0.071)	-0.036 (0.070)
Treatment 1 x Female	0.011 (0.098)	0.033 (0.100)
Controls	No	Yes
SE Type	Robust	Robust
Observations	425	425

*Notes:* Table 13 shows the outcome for whether or not a student is interested in receiving networking emails on gender. Entries are OLS coefficients. Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 14: OLS Estimation – Networking Email for Female Students

	(1) Networking Emails	(2) Networking Emails
Treatment 2	-0.046 (0.080)	-0.062 (0.079)
Female	-0.013 (0.071)	-0.032 (0.072)
Treatment 2 x Female	0.086 (0.099)	0.097 (0.098)
Controls	No	Yes
SE Type	Robust	Robust
Observations	433	433

*Notes:* Table 14 shows the outcome for whether or not a student is interested in receiving networking emails on gender. Entries are OLS coefficients (risk differences). Robust standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

# A Appendix

## A.1 Additional Tables and Figures

Table A.1: Means by First Generation Status

	First-Gen		Continuing-Gen		Diff
	Mean	(SD)	Mean	(SD)	
Cumulative GPA	3.20	( 0.63)	3.43	( 0.55)	-0.23***
Credit-Bearing Internship	0.17	( 0.37)	0.20	( 0.40)	-0.03***
Non-Credit Internship	0.17	( 0.38)	0.26	( 0.44)	-0.09***
Observations	19,126		97,027		

*Notes:* Table A.1 shows means, standard deviations (in parentheses), and differences across first-generation and continuing-generation students. Stars indicate significance levels: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .



Table A.2: Current GPA with HS GPA as Control

	Cumulative GPA	Cumulative GPA
	(1)	(2)
First-Generation	-0.125*** (0.006)	-0.069*** (0.007)
High School GPA	0.573*** (0.010)	0.577*** (0.011)
Gender	0.012*** (0.004)	0.022*** (0.004)
Hispanic	-0.259*** (0.007)	-0.145*** (0.009)
American Indian	-0.056 (0.048)	-0.017 (0.054)
Asian	0.005 (0.006)	0.071*** (0.008)
Black	-0.247*** (0.015)	-0.146*** (0.017)
Pacific Islander	-0.150*** (0.048)	-0.090* (0.050)
White	-0.035*** (0.006)	0.012 (0.008)
Low Income		0.000 (.)
Working Class		0.032*** (0.008)
Middle Class		0.134*** (0.008)
Upper Middle Class		0.202*** (0.009)
Wealthy		0.220*** (0.014)
In State		-0.018** (0.007)
US Citizen		-0.110*** (0.013)
Observations	59,644	48,520
Controls 1	Yes	Yes
Controls 2	Yes	Yes
University FE	Yes	Yes

*Notes:* Table A.2 shows results from an OLS regression with University level fixed effects. Stars indicate significance levels: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table A.3: First-Gen and High School GPA

	High School GPA		
	(1)	(2)	(3)
First-Generation	-0.037*** (0.003)	-0.026*** (0.003)	-0.019*** (0.004)
Observations	62,386	61,707	50,412
Controls 1	No	Yes	Yes
Controls 2	No	No	Yes
University FE	Yes	Yes	Yes

*Notes:* shows results from OLS regression on high school GPA with university level fixed effects. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels:  $*p < 0.1$ ,  $**p < 0.05$ ,  $***p < 0.01$ .

Table A.4: First Generation and No Credit Internships

	No Credit Internships	
	(1)	(2)
First-Generation	-0.093*** (0.008)	-0.092*** (0.009)
Observations	72,189	58,925
Controls 1	Yes	Yes
Controls 2	No	Yes
University FE	Yes	Yes

*Notes:* Table A.4 shows results from a Probit regression on whether student did a for no credit internship. Entries are average marginal effects (AME) on the probability scale. Standard errors in parentheses.  $* p < 0.10$ ,  $** p < 0.05$ ,  $*** p < 0.01$ . For binary regressors, AME is the discrete change from 0 to 1.

Table A.5:  $R^2$  Table on For Credit Internships

	Credit Internships					
	(1)	(2)	(3)	(4)	(5)	(6)
# Professors	0.048*** (0.001)	0.047*** (0.001)	0.048*** (0.001)		0.067*** (0.001)	0.044*** (0.001)
Student Org	0.064*** (0.004)	0.068*** (0.004)	0.069*** (0.004)		0.069*** (0.004)	0.069*** (0.003)
Entrepreneurial	0.043*** (0.011)	0.037*** (0.010)	0.042*** (0.010)		0.065*** (0.011)	0.045*** (0.008)
Leadership	0.075*** (0.004)	0.076*** (0.004)	0.073*** (0.004)		0.095*** (0.004)	0.075*** (0.003)
Class Presentations	0.016*** (0.001)	0.016*** (0.001)	0.015*** (0.001)	0.031*** (0.001)		0.017*** (0.001)
Readings	-0.010*** (0.001)	-0.009*** (0.001)	-0.010*** (0.001)	-0.008*** (0.001)		-0.010*** (0.001)
Methods Class	0.057*** (0.003)	0.060*** (0.003)	0.058*** (0.003)	0.076*** (0.004)		0.057*** (0.003)
Independent Research	0.116*** (0.004)	0.117*** (0.004)	0.112*** (0.004)	0.145*** (0.004)		0.111*** (0.003)
High School GPA	0.001 (0.002)	0.002 (0.002)	0.002 (0.002)	0.003 (0.003)	-0.000 (0.002)	
Observations	49,090	49,613	54,014	49,356	49,737	93,745
Pseudo R-sq.	0.131	0.128	0.130	0.091	0.094	0.116
Controls 1	Yes	No	Yes	Yes	Yes	Yes
Controls 2	Yes	Yes	No	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* Table A.5 results from a Probit regression on for credit internship with university level fixed effects. Controls 1 are: gender and ethnicity. Controls 2 are: social class dummy, being from in-state, and dummy for U.S. nationality. Stars indicate significance levels: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table A.6: OLS Estimation – Networking Email for Transfer Students

	(1) Networking Emails	(2) Networking Emails
Treatment 1	-0.053 (0.052)	-0.047 (0.053)
Transfer	-0.207*** (0.076)	-0.197** (0.078)
Treatment 1 x Transfer	0.125 (0.111)	0.098 (0.113)
Controls	No	Yes
SE Type	Robust	Robust
Observations	425	425

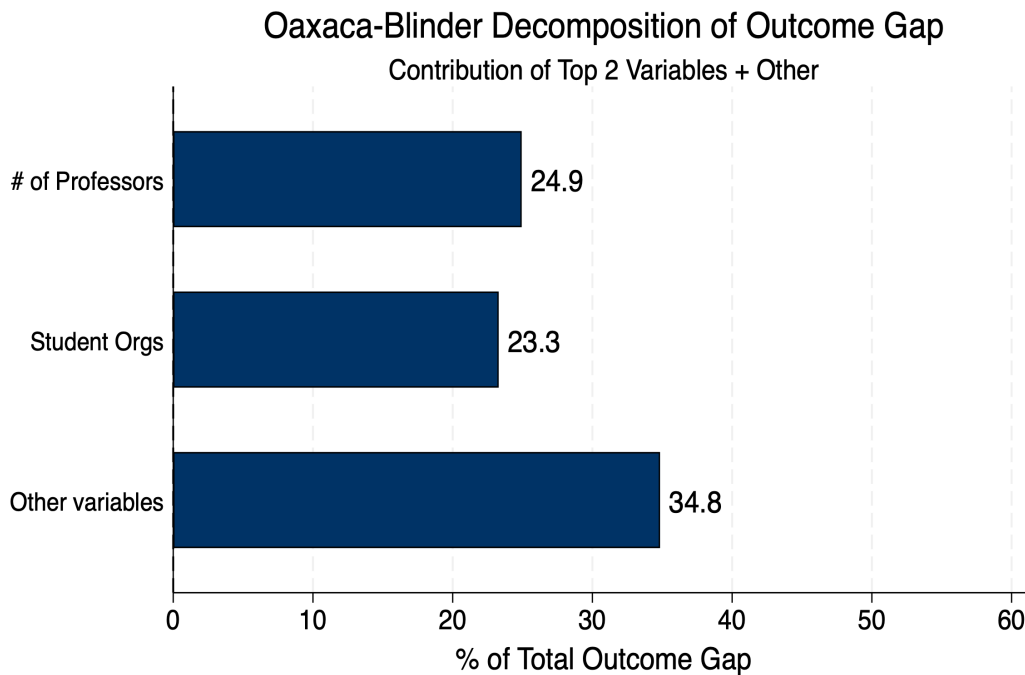
*Notes:* Table A.6 shows the outcome for whether or not a student is interested in receiving networking emails on community college transfer status. Entries are OLS coefficients (risk differences). Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table A.7: OLS Estimation – Networking Email for Transfer Students

	(1) Networking Emails	(2) Networking Emails
Treatment 2	-0.050 (0.053)	-0.057 (0.053)
Transfer	-0.207*** (0.076)	-0.197** (0.080)
Treatment 2 x Transfer	0.310*** (0.110)	0.307*** (0.113)
Controls	No	Yes
SE Type	Robust	Robust
Observations	433	433

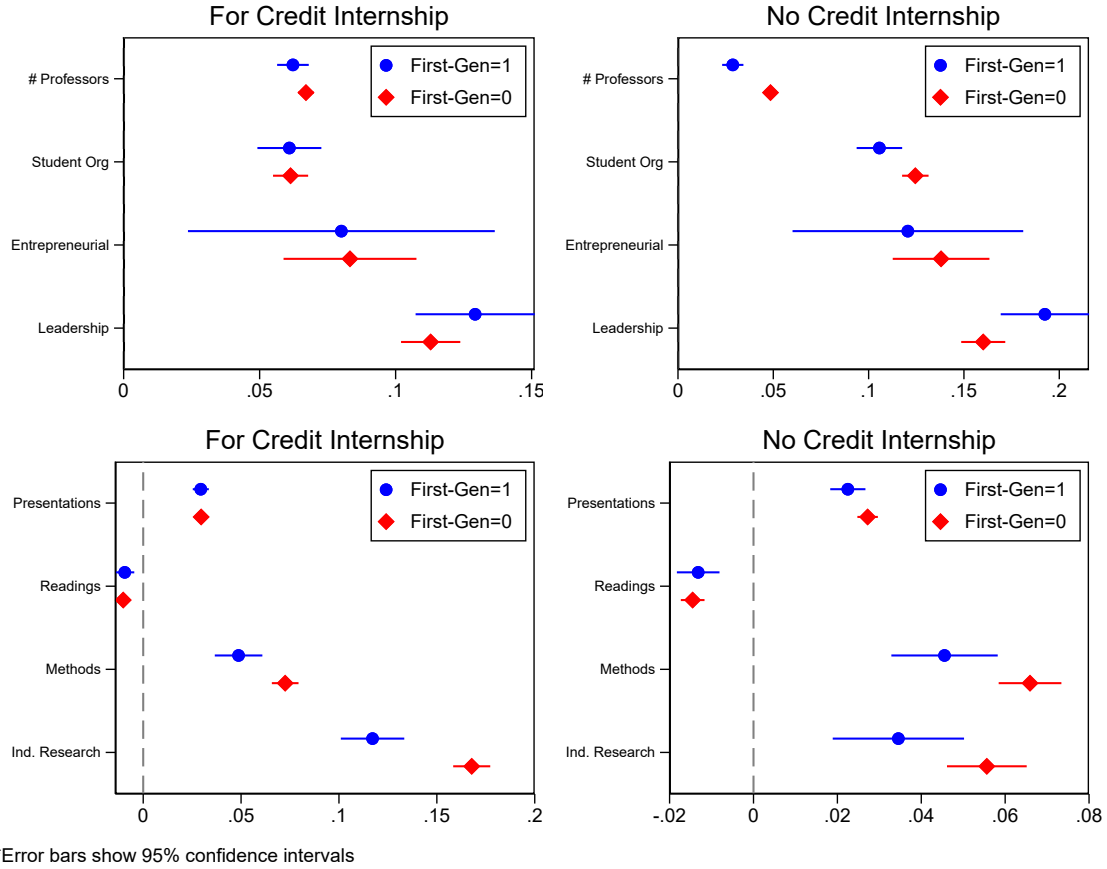
*Notes:* Table A.7 shows the outcome for whether or not a student is interested in receiving networking emails on community college transfer status. Entries are OLS coefficients (risk differences). Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure A.1: Oaxaca-Blinder Decomposition of Internships



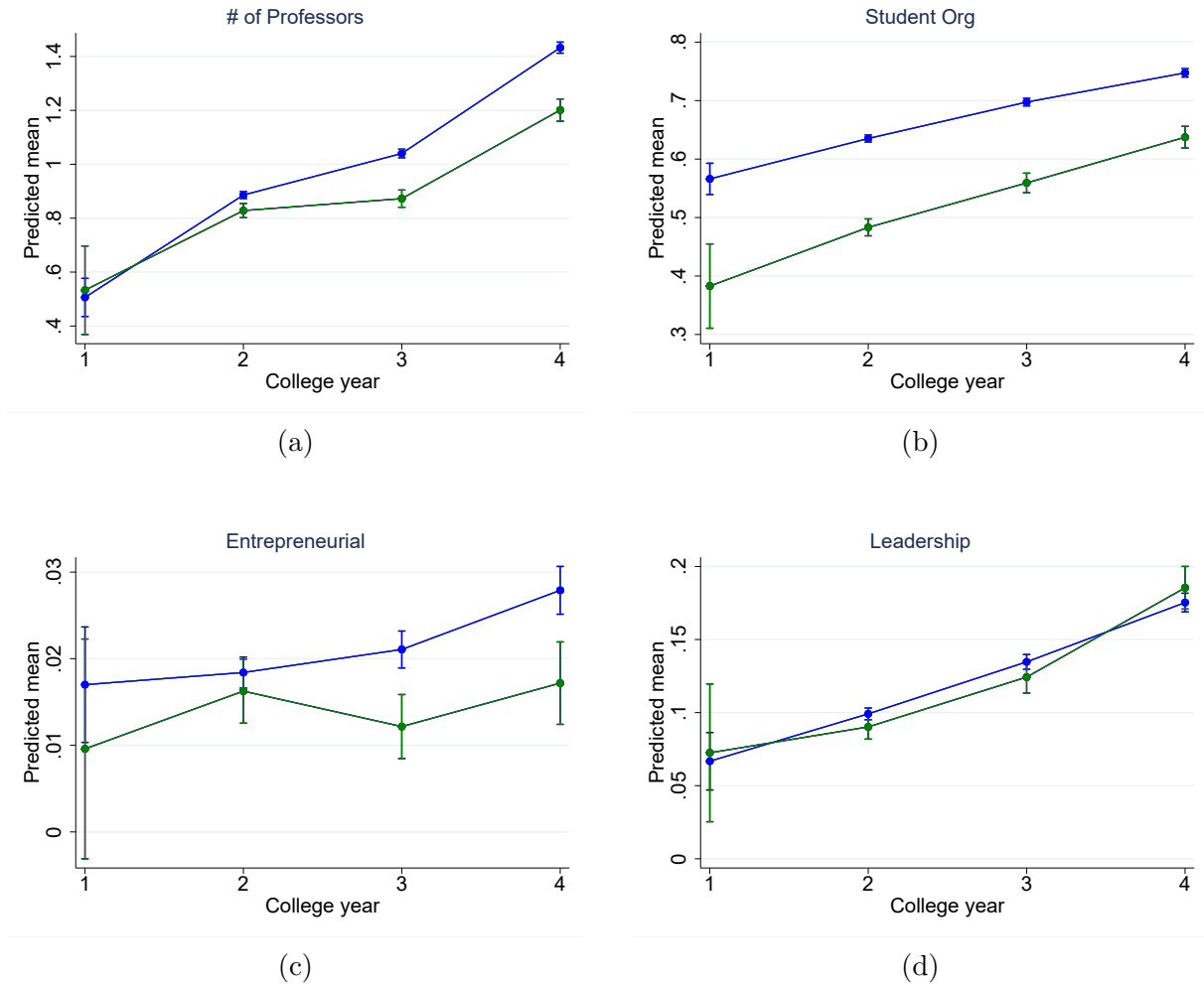
*Notes:* The figure shows contribution of top 2 variables and all others to the outcome gap between first-generation and continuing-generation students in Internships. Contribution was determined from a Oaxaca-Blinder decomposition. Number of faculty and joining a student organization are the top two variables and explain 48.2% of the gap. In other variables we have: gender dummy, ethnicity dummies, social class dummies, same state as university dummy, nationality dummy, entrepreneurial activity dummy, leadership program dummy, in class presentations, class readings, methods class, independent research activity, high school GPA, and university dummy.

Figure A.2: Coefficient Comparison by First-Gen Status



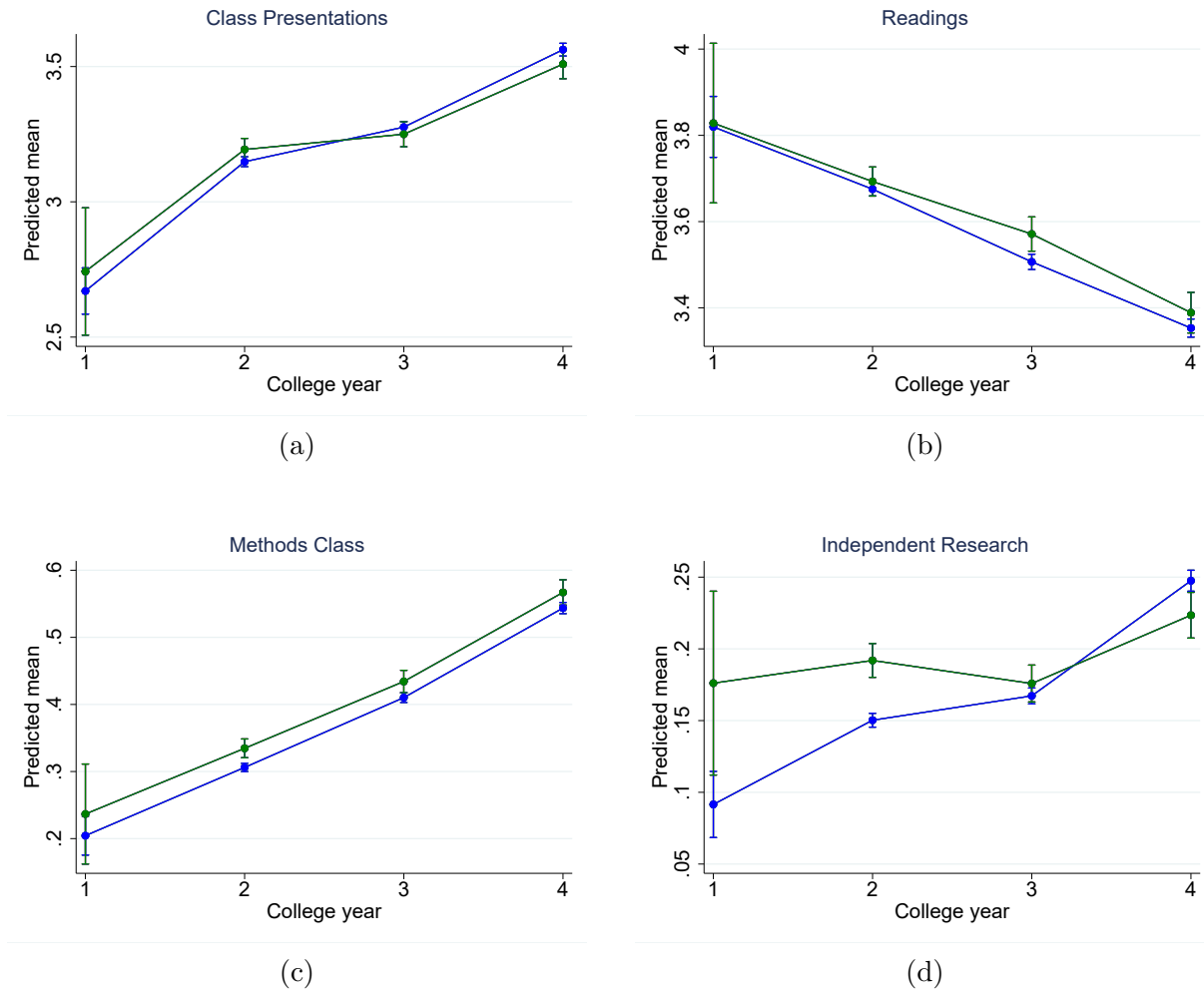
*Notes:* The figure shows a OLS regression on observables and reports the coefficients for actions. On the two left panels the outcome variable is whether the student secures a for credit internship. The two right panels are for no credit internship. The top row is for hidden curriculum actions: number of faculty they network with, student organizations, entrepreneurial activity, and leadership programs. The bottom row is for formal curriculum: class presentations, class readings, methods class, and independent research. Regressions include the following controls: gender dummy, ethnicity dummies, social class dummies, same state as university dummy, nationality dummy. Regressions include university fixed effects. Bars indicate 95% confidence intervals.

Figure A.3: Differences in Hidden Curriculum Levels



*Notes:* The figure shows levels of HC activities for first-generation and continuing-generation students. The bars represent the 95% CI. In Panel A.3a the outcome variable is the number of Professors the student networks with. In Panel A.3b the outcome variable is joining a student organization. In Panel A.3c the outcome is official entrepreneurial activities. In Panel A.3d the outcome variable is participating in leadership programs.

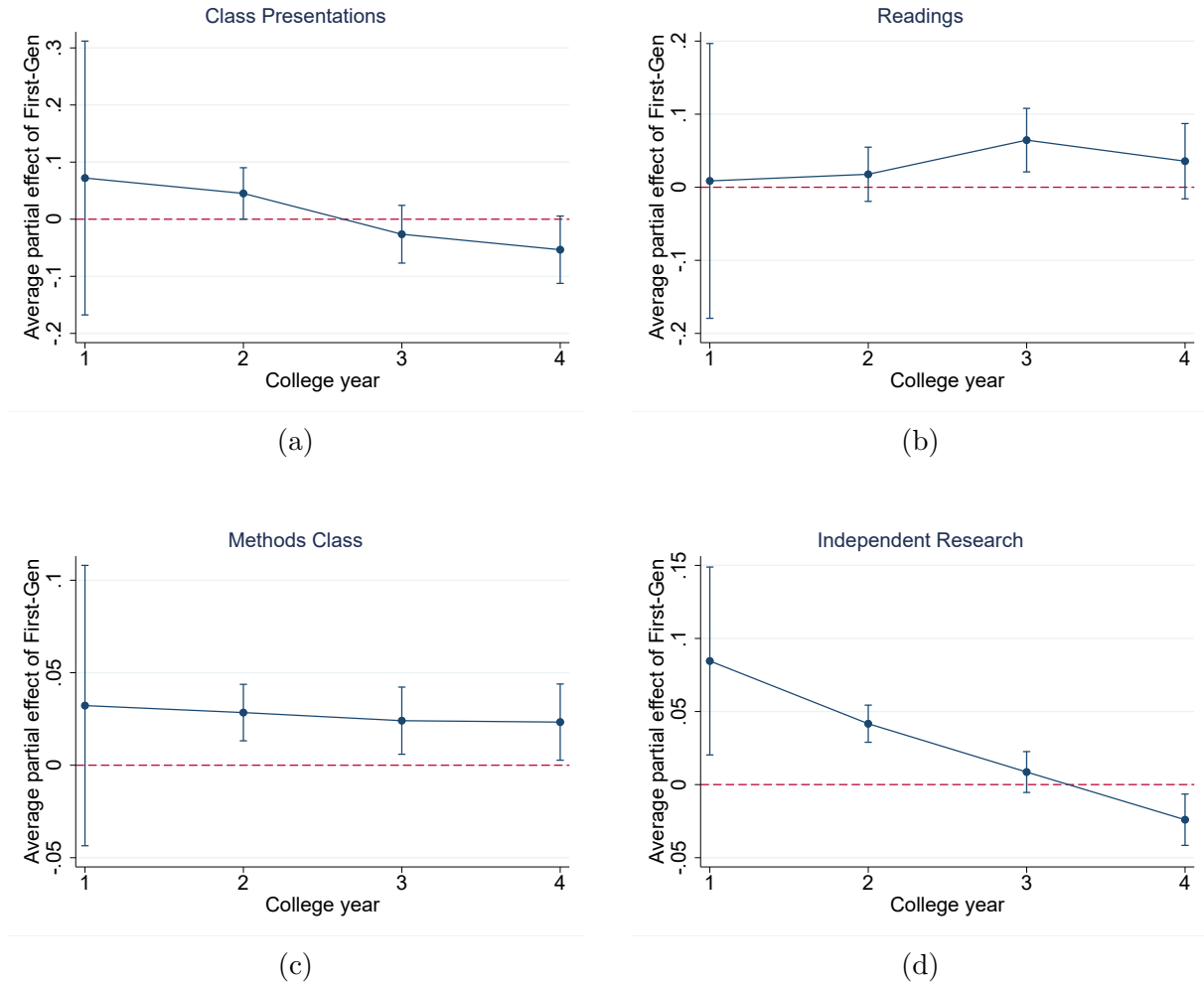
Figure A.4: Differences in Formal Curriculum Levels



*Notes:* The figure shows levels of formal curriculum activities for first-generation and continuing-generation students. The bars represent the 95% CI. In Panel A.4a the outcome variable is class presentations. In Panel A.4b class readings. In Panel A.4c the outcome is doing a methods class. In Panel A.4d the outcome variable is independent research.



Figure A.5: Differences in Formal Curriculum Levels



*Notes:* The figure shows differences of formal curriculum activities for first-generation and continuing-generation students. The bars represent the 95% CI. In Panel A.5a the outcome variable is class presentations. In Panel A.5b class readings. In Panel A.5c the outcome is doing a methods class. In Panel A.5d the outcome variable is independent research.

## A.2 Additional Outcomes

In this section we report additional outcome measures from our observational dataset. In particular we look at more qualitative items that represent students perspectives and experiences on campus.

We find that overall first-generation students feel no different in their sense of belonging, their academic satisfaction, social life, and in their willingness to re-enroll in the same university. However interestingly enough first-generation students, as reported in Table A.8, are more likely to report feeling respected and that their education is valuable.

Table A.8: Additional Outcomes

	(1) Feel Respected	(2) Feel Belong	(3) Feel Satisfied Academically	(4) Feel Satisfied Socially	(5) Satisfied Value of Edu	(6) Enroll Again
First-Generation	0.080*** (0.012)	-0.001 (0.012)	0.011 (0.010)	-0.004 (0.012)	0.083*** (0.013)	0.016 (0.013)
Observations	96,236	96,144	96,126	95,811	95,779	96,149
Controls 1	Yes	Yes	Yes	Yes	Yes	Yes
Controls 2	Yes	Yes	Yes	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* Table A.8 Estimates from an OLS regression. Robust standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Controls 1 include gender and race/ethnicity; Controls 2 include social class and geography.

## B Model

We define  $h \geq 0$  as the investment in **hidden curriculum** activities and  $v \geq 0$  as the investment in the **formal curriculum**. Ability  $a \geq 0$  is exogenous and identical across groups in the comparative statics below.

**Technology** The quality index is

$$Q = f(h, v, a), \quad (\text{A.1})$$

with  $f : \mathbb{R}_+^3 \rightarrow \mathbb{R}_+$  twice continuously differentiable, strictly increasing and concave in  $(h, v, a)$ . Partial derivatives are denoted as  $f_h, f_v, f_{hh}, f_{vv}, f_{hv}$ .

**Wages** Wages<sup>33</sup> are  $W : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  twice continuously differentiable, strictly increasing and concave:  $W'(Q) > 0$ ,  $W''(Q) \leq 0$ .

**Costs**  $C : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$  is twice continuously differentiable, increasing and convex:  $C_h, C_v > 0$ ,  $C_{hh}, C_{vv} \geq 0$ . In several results we assume separable costs,  $C(h, v) = c_h(h) + c_v(v)$  (implying  $C_{hv} = 0$ ).

**Objective.** A student of type  $t \in \{\text{CG}, \text{FG}\}$  chooses  $(h, v)$  to

$$\max_{h, v \geq 0} U^t(h, v) \equiv [W\{\hat{f}^t(h, v, a)\}] - C(h, v), \quad (\text{A.2})$$

where the expectation captures shocks orthogonal to  $Q$  (no uncertainty about  $W$  conditional on  $Q$ ). In the benchmark, CG has correct beliefs<sup>34</sup>  $\hat{f}^{\text{CG}} \equiv f$ . FG may face one of two informational frictions:

1. **Awareness:**  $h$  not in the consideration set.
2. **Subjective returns:**  $\hat{f}^{\text{FG}} \neq f$  with a uniformly lower marginal value of  $h$ :

$$\left[ W'(\hat{f}^{\text{FG}}(h, v, a)) \hat{f}_h^{\text{FG}}(h, v, a) \right] \leq \left[ W'(f(h, v, a)) f_h(h, v, a) \right] \quad \forall (h, v). \quad (\text{A.3})$$

A convenient special case is

$$\hat{f}_h^{\text{FG}}(h, v, a) = \lambda f_h(h, v, a), \quad \lambda \in (0, 1], \quad (\text{A.4})$$

---

<sup>33</sup>Wages can be interpreted as lifetime utility from wages or monetary outcomes.

<sup>34</sup>This is not strictly needed for results on FG students.

with  $\hat{f}_v^{\text{FG}} = f_v$ .

## B.1 Existence, Uniqueness, and First-Order Conditions

Define the “true” objective  $\Phi(h, v) \equiv [W(f(h, v, a))] - C(h, v)$ . By the assumptions on  $W$  and  $f$ ,  $[W(f(\cdot))]$  is concave in  $(h, v)$ ; subtracting the convex  $C$  keeps  $\Phi$  concave. If either  $W \circ f$  is strictly concave or  $C$  strictly convex, the optimizer  $(h^*, v^*)$  is unique.

For interior optima the FOCs are

$$[W'(Q) f_h(h, v, a)] = C_h(h, v), \quad [W'(Q) f_v(h, v, a)] = C_v(h, v). \quad (\text{A.5})$$

### Two useful reduced forms.

1. *Value in  $h$ .* Let  $v^*(h) \in \arg \max_{v \geq 0} \Phi(h, v)$  and  $V(h) = \Phi(h, v^*(h))$ . The envelope theorem gives

$$V'(h) = [W'(Q) f_h(h, v^*(h), a)] - C_h(h, v^*(h)). \quad (\text{A.6})$$

$V$  is concave in  $h$ .

2. *Reaction of  $v$  to changes in  $h$ .* Totally differentiate the second condition in (A.5) along the optimal locus, obtaining

$$\frac{v}{h} = \frac{[W'' f_h f_v + W' f_{hv}] - C_{hv}}{C_{vv} - [W'' f_v^2 + W' f_{vv}]}, \quad (\text{A.7})$$

where the denominator is positive by the (negative-definite) SOC.

## B.2 Proof of Proposition 1

For students with identical ability  $a$ , the equilibrium  $h$  satisfies  $h_{\text{FG}}^* \leq h_{\text{CG}}^*$  if either (i)  $h$  is absent from the FG consideration set, or (ii) FG holds lower marginal-return beliefs about  $h$  as in (A.3) (e.g., the wedge (A.4) with  $\lambda < 1$ ).

(i) (*Awareness*). If  $h$  is not in the FG consideration set, feasibility imposes  $h = 0$ . Since the CG optimum has  $h_{\text{CG}}^* \geq 0$ , we have  $h_{\text{FG}}^* = 0 \leq h_{\text{CG}}^*$ .

(ii) (*Subjective returns*). Consider  $V(h) = \max_v \Phi(h, v)$  for CG and  $\tilde{V}(h) = \max_v \tilde{\Phi}(h, v)$  for FG, where  $\tilde{\Phi}(h, v) = [W(f^{\text{FG}}(h, v, a))] - C(h, v)$ . By (A.6),

$$V'(h) = [W'(f) f_h] - C_h, \quad \tilde{V}'(h) = [W'(f^{\text{FG}}) f_h^{\text{FG}}] - C_h.$$

By (A.3),  $\tilde{V}'(h) \leq V'(h) \forall h$ . Let  $h_{\text{CG}}^*$  solve  $V'(h) = 0$ . Then  $\tilde{V}'(h_{\text{CG}}^*) \leq 0$ . As  $\tilde{V}$  is concave, the root of  $\tilde{V}'(h) = 0$  lies weakly to the left:  $h_{\text{FG}}^* \leq h_{\text{CG}}^*$ .

In the multiplicative wedge case (A.4),  $\tilde{V}'(h) = \lambda [W'(f)f_h] - C_h$ ; at  $h_{\text{CG}}^*$  this equals  $(\lambda - 1)C_h < 0$ , yielding a strict inequality.

### B.3 Proof of Proposition 2

Let  $(h^*, v^*)$  satisfy (A.5). The sign of the equilibrium response of  $v$  to changes in  $h$  is

$$\text{sign}\left(\frac{v}{h}\right) = \text{sign}\left(\underbrace{[W''f_hf_v]}_{\leq 0} + \underbrace{[W'f_{hv}]}_{\text{production complementarity}} - \underbrace{C_{hv}}_{\text{cost complementarity}}\right). \quad (\text{A.8})$$

With separable costs ( $C_{hv} = 0$ ): if  $f_{hv} \leq 0$  (substitutes or independence), then  $v/h \leq 0$ ; if  $f_{hv} > 0$  (complements), the sign is ambiguous because complementarity (+) must overcome the negative “income” term  $[W''f_hf_v]$ .

From (A.7),  $\text{sign}(v/h)$  is the sign of the numerator. Because  $W'' \leq 0$  and  $f_h, f_v > 0$ , the first term is non-positive; the remaining terms capture production and cost complementarity.

**Interpretation.** Starting from the CG optimum where  $h$  is higher than under FG (Prop. 1), (A.8) implies:

- **Substitutes/independence** ( $f_{hv} \leq 0$ ):  $v$  must be lower at higher  $h$ . Hence  $v_{\text{FG}}^* \geq v_{\text{CG}}^*$ .
- **Complements** ( $f_{hv} > 0$ ): the direction is a priori undetermined; strong complementarity can overturn the (negative) income term and yield  $v_{\text{FG}}^* < v_{\text{CG}}^*$ .

### B.4 Multi-Dimensional Extension (sketch)

Let  $h \in_+^k$  and  $v \in_+^m$ . With the same primitives and strict concavity, all results generalize. Proposition 1 follows by applying the envelope argument to the reduced value  $V(h) = \max_v \Phi(h, v)$  and a vector version of (A.3). Proposition 2 extends componentwise: for each formal input  $v_j$ ,

$$\frac{\partial v_j}{\partial h_\ell} = \frac{[W''f_{h_\ell}f_{v_j}] + [W'f_{h_\ell v_j}] - C_{h_\ell v_j}}{C_{v_j v_j} - [W''f_{v_j}^2 + W'f_{v_j v_j}]},$$

with the same “income vs. complementarity” interpretation as in (A.8).

## C Information Experiment Elicitations

Below we report the two information treatment interventions.

*We have talked to many hiring managers, alumni, and other professionals in the field. We asked them what they thought where invaluable information that most students are not aware of.*

**T2** *These are the points all of them agreed on: You do not need to know someone directly in order to network! Common links, like being from the same undergraduate institute, help a lot when contacting people. You do not need to rely on personal connections. Networking is not only a great way to learn more about the industries and firms you are interested in, but also to discover new industries and firms. Talking to alumni also sends a message that you are putting in the effort to learn about the firm, even if you donât know anything about it right now. Hiring managers have told us that “Engaging with alumni can give you a competitive edge and help you get noticed in a crowded job market” and that “The best candidates often come from referrals”. Alumni say about prospective applicants “they need a connection....it doesn’t need to be their immediate family member, or friend, or whatever, but they need something”. Not a senior? Not looking for a job yet? It doesnât matter! Itâs never too early to start learning about jobs. Actually early internships, even if not in your exact desired role, can help a lot in getting noticed down the line. When we surveyed hiring managers, they emphasized that almost no one in \$career/ChoiceGroup/SelectedChoices can secure a job without first making connections within the firm.*

**T1** *We are trying to learn what students value the most in preparation for a job in your career of interest. For example we are interested in seeing whether some available and accepted strategies like networking are underutilized.*

## D AI Instruction

### D.1 Active AI

You are a college career advisor assistant. Your job is to help students navigate their college experience in order to pursue career paths that match their interests and goals.

You are in the treatment condition of an experiment. Your goal is to get the student to consider and for them to take concrete steps in the following 6 topics :

- Networking (peers/alumni/cold outreach)
- Office hours (strategic use beyond homework help)
- Relationship building with professors
- Professional organizations/extracurriculars
- Research opportunities
- Mentorship seeking

If the student does not ask about these topics, you should suggest they think about the topics. Ask if they are taking these important actions and if they have any questions about how to get started on these actions.

If they continue to ignore some of the actions, explain why these other actions are important. You should use the information below as part of your explanation<sup>35</sup>.

## D.2 Passive AI

You are a college success Q&A assistant participating in a research study. You must follow these rules STRICTLY: CORE RULES:

- ONLY answer the specific question asked - never volunteer additional information
- Never mention topics the user hasn't explicitly asked about
- If asked a vague question, request clarification rather than providing broad information
- Keep answers concise (2-3 sentences max unless asked for more detail)
- Never suggest related topics or say things like "you might also want to know about..."

---

<sup>35</sup>Here information below is referring to JSONs built to make sure the hidden curriculum topics brought up were appropriate to the students' needs.