PhD-FDEF-2025-033
The Faculty of Law, Economics and Finance

DISSERTATION

Defence held on 12/12/2025 in Esch-sur-Alzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN *SCIENCES ECONOMIQUES*

by

Reha TUNCER
Born on 6 July 1996 in Rotterdam, The Netherlands

Essays in Experimental Economics: Digital Nudging and Social Capital

Dissertation defence committee

Dr Ernesto Reuben, dissertation supervisor
Professor, New York University Abu Dhabi

Dr Christine Schiltz, Chairman
Professor, Université du Luxembourg

Dr Manuel Muñoz Herrera
Research Scientist, New York University Abu Dhabi

Dr Moritz Janas
Professor, University of Gothenburg

Dr Bertrand Verheyden
Research Scientist, Luxembourg Institute of Socio-Economic Research

Contents

List of Figures

List of Tables

Acknowledgements

I gladly accepted the offer for this PhD project almost exactly four years ago. Little did I know life would test me so hard, harder than the econometrics exams I had passed to come to this point. I am thankful to myself and those around me for not letting me give up. Perseverance, it seems, is my defining personality trait. My seeing the light at the end of this academic tunnel is not my own doing alone but a collective effort—something you don't learn in econ grad school or by being an only child.[1]

First and foremost, my dear Teresa. I sincerely do not know how you managed to keep it together and keep us together. You saw how this experience changed me, and I hope to change back to the person you fell in love with. I am grateful that you are always by my side.

Next, the second victim of my dissertation would be Mamo, my closest friend. You are definitely the person with whom I spend the most time after my fiancée. I am grateful for every single game night we shared and for the value you place on scientific work, especially in moments when I tended to forget it.

Then my parents, who I am sure have always happily listened to my complaining over the phone. I am grateful that you are watching over me as I surpass myself with this endeavor. Thank you for helping me become the person I am.

I am grateful for my two elder 'brothers'. Çağrı sparked my interest in behavioral economics by handing me the famous *Thinking, Fast and Slow* by Kahneman. Okan's interest in automating mundane processes with Python while brainstorming with me around a

[1] We will touch upon the 'social' in Chapters 2 and 3 thanks to the direction the thesis took after year two.

glass of whiskey ultimately led to my first publication in economics. There is no shame in saying how you both have profoundly shaped how I see the world and my place in it.

I am also grateful for the friends I made along the way. Suvadeep, Amaury, Xiaowei, Roxane, Laura, Stella, Nathalie, Gesine, Alper, Margault and the 3E cohort made this journey enjoyable, helping me see there is more to office life than work. I am thankful to the CNS Lab members with whom I shared my office, for always inviting me for lunch and their activities, even if we have completely different research interests. To my senior co-authors Kerstin, Anastasia, Sophie and Jhon – I am grateful to have had the opportunity to collaborate with you, and thank you for all your feedback.

Finally, I am grateful for my supervisors Ernesto, Manu, and Christine, whose guidance and support were simply indispensable through these years. I am lucky to have been able to learn from your experience. Thank you for your kindness and patience.

Abstract

This dissertation consists of three experimental studies documenting how structural constrains alter economic decisions. The constraints I study are digital in chapter one, and become institutional later. In chapter one, I study the effects of a specific interface design feature called autoplay from the perspective of digital nudging and temptation. Using the intertemporal choice framework, I test whether autoplay causes preference reversals and increases content consumption. I find that autoplay does not cause an increase in content consumption compared to the control condition, and that participants are willing to pay a small positive sum to have autoplay videos. My results suggest the experimental environment itself influenced participant behavior, driving down their content consumption and masking the true effects of autoplay. I conclude by underscoring the need to assess digital nudges in the field, where decisions take place more naturally. The opportunity to collect data directly from the field appears in the later chapters, with the subject of my inquiries changing to institutional factors influencing social capital. In the following chapter, I study how classroom interactions at a university shapes learning about peer skills and transmission of opportunities in the form of referrals between classmates. I ask whether cognitive and social skill signals can be accurately transmitted in the classroom, and whether referrals can flow to disadvantaged peers by randomly assigning participants within the same classroom to receive additional incentives. I find that classroom interaction during a semester results in learning about cognitive but not social skill, and that the treatment with additional incentives to refer disadvantaged peers mitigate biases without compromising performance. Inspired by these positive results, in the final chapter, I study university-wide referral networks. I randomly assign participants to a treatment where on top of receiving earnings based on the performance of their referral, the candidate they pick gets a sizeable fixed monetary bonus. I find that university-wide referrals go

to higher performing students with whom the referrer has taken many courses together, regardless of the treatment. Referrers are also not biased in their individual preferences against picking disadvantaged candidates. Yet, this lack of bias does not mean referrals flow from wealthier to disadvantaged peers. I find that student networks are segregated by SES, especially in the parts where referrals emerge from, with academic program selection as a key driver of this segregation. With program fees determined by their cost, and a lack of scholarships, disadvantaged students sort into affordable programs. They lack the opportunities to meet wealthier peers except in rare courses which were studied in chapter two. In sum, I find that institutional factors have a crucial role in the accumulation of social capital. As the results from this dissertation shows the beneficial impact of intergroup contact, future research should look into ways of increasing the exposure of disadvantaged students to their wealthier peers.

Use of AI disclaimer

During the preparation of the thesis, the author used Claude Sonnet 3.5 (2024) from Anthropic, for proofreading and to ensure linguistic precision. After using the tool, the author reviewed and edited the text as needed, and takes full responsibility for the content and wording of the thesis.

Chapter 1

Introduction

> *"Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist."*
>
> – Herbert A. Simon

The overall purpose of this thesis is to study and illustrate examples of economic behavior constrained by structural factors which influence people's choices. It is motivated by the long line of literature documenting the ways in which individuals diverge from the *homo economicus* ideal, a decision-maker who always chooses the best consumption bundle among all affordable options (Thaler, 2017). In this introduction, I will demonstrate how the three chapters link together, and provide relevant theoretical and methodological background behind the creation of each chapter. I will define and explain concepts as I introduce them, and also provide a discussion based on my findings.

The first over encompassing theme in this dissertation is digital nudging in the domain of consumer products (Wendel, 2016), and how the online decision environment shapes behavior. Within this framework, all firms that present people with choices are choice architects (Johnson et al., 2012), devising digital interventions including nudges. "[A nudge] is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2009, p.8). In this sense, the first goal of this dissertation is studying how isolated nudges in the online decision environment could alter behavior – a non-controversial extension of behavioral economics literature.

My focus is solely on private sector nudges that can lead to undesirable outcomes for consumers. The inspiration comes from the idea that "[free markets] create an economic equilibrium that is highly suitable for economic enterprises that manipulate or distort our judgment [...] Insofar as we have any weakness in knowing what we really want, and also insofar as such a weakness can be profitably generated and primed, markets will seize the opportunity to take us in on those weaknesses." (Akerlof & Shiller, 2015, p. x). As Akerlof and Shiller neatly put it, once we accept the plethora of evidence collected in the last half of the twentieth century on the cognitive limitations of human decision-makers (Hanson & Kysar, 1999), we must also accept that in the absence of regulation, firms have inherent competitive pressures to profit from the same limitations.

The experimental study of these nudges within the online choice environment defines the conceptual scope of the first part of my dissertation. This first theme is self-contained within chapter one, in the sense that I consecrate a single experiment to this framework. Chapter one is an experiment about a digital nudge introduced by private firms accused of tempting people to alter their streaming/social media content consumption.

The second over encompassing theme in this dissertation concerns the sharing of job-market relevant information within social networks (Beaman, 2016; Jackson, 2008). Consisting of a broad spectrum of interactions within networks that facilitate job matches between firms and workers, Topa (2019) suggests "[t]hese interactions range from the simple transmission of information about job openings at a particular firm to the provision of a referral: namely, recommending a social contact to a potential employer to be hired for a given position." They lie on a spectrum, involving different information content and social costs. On one end, letting someone know about a job involves only knowledge of the existence of the job opening, no communication between the referrer and the employer, and very little reputational cost for the referrer. At the other end, recommending someone to an employer for a given position involves explaining why the referrer thinks their referral would be a good match for the employer, and a large reputational cost for the referrer if the new match does not work out as expected. In this sense, the second goal of this dissertation is the study of how labor-market relevant information inside social networks propagate.

I focus on the role of social capital in these interactions. According to Lin (2008), 'capital' here describes the investment in valuable resources (as in the role of investment in education for nurturing human capital), and refers to the process by which it is captured and reproduced for returns. Social capital in this sense can be defined in two ways. First, it can be understood as the potential access to resources embedded in a network. This potential is quantified by "the size of the network of connections he can effectively mobilize and on the volume of the capital possessed in his own right by each of those to whom he is connected." (Bourdieu, 1986, p. 251). Second, social capital can also be defined in terms of its actual use, or 'mobilization', which involves selecting specific connections and their resources (like wealth or status) for a particular goal, such as a job search. I refer to social capital both from the perspective of one's access to it and one's mobilization of it. Either way, social capital is contingent on social networks, with "networks providing the necessary condition for access to and use of embedded resources." (Lin, 2008, p. 58).

My research interest in this part of the dissertation therefore lies in one's capacity and use of their connections for the labor market. I build on the fact that having connections with a higher share of wealthier individuals in one's network strongly correlates with higher labor market earnings (Chetty et al., 2022). This piece of evidence, aligning with the theory presented earlier, suggests that increasing the access to the pool of embedded resources (in the form of the wealth of one's connections) within the network is instrumental in generating measurable returns for the network owners.

Given an exogenous access to a pool of resources inside a network, I am interested in how the mobilization of social capital, in the form of referrals, can be utilized to access a bigger pool of embedded resources. After observing the distribution of opportunities (referrals) in networks and the initial differences in social capital access across individuals, I experimentally study whether these differences can be remedied by incentives in the mobilization process. In an experimental study in chapter two, I consider skill-based referrals in a classroom setting, where classmates learn about each others' skills and group identities. In chapter three I extend my reach to the entire university networks of the students and experimentally study how incentives impact university-wide referral decisions. I provide a close inspection of the institutional constraints that limit interactions across different student groups and sharing of valuable opportunities.

To sum, each of the three chapters thus characterizes situations where decision makers are constrained in some palpable way. In chapter one, I consider how firms can nudge people to change digital behavior; in chapter two I consider learning about and referring classmates when signals about skills are blurry. In chapter three, I look at how incentives within an institution curb interactions across groups and the sharing of valuable opportunities. Throughout, I focus on the economic consequences of the constraints, either from the consumers of digital content (streaming and social media) perspective in the first chapter, or the labor market consequences for hiring using social connections (referrals) in the last two chapters.

1.1 Temptation and digital nudges

1.1.1 Theoretical background and motivation

The constraints I study are digital in chapter one. I draw heavily from dark patterns literature,[1] where user interface designers identify questionable practices in the digital choice environment that go against users' best interests. The academic work on the topic began with various classification attempts based on existing design practices in the web (e.g., cookie consent banners, default/preselected options) and focused on their malicious intent while changing behavior (Bösch et al., 2016; Gray et al., 2018). While proving the malicious intent behind the implementation of design features is an elusive task,

[1] The origin of the term 'dark pattern' comes from Harry Brignull's website where an initial body of common industry practices were collected.

meaning we can never know for sure why a firm chooses to implement a design feature over another, a more convenient normative perspective is focusing on their impact on individual welfare. From this perspective, "a dark pattern is any interface that modifies the choice architecture to benefit the designer at the expense of the user's welfare." (Mathur et al., 2021, p. 9).

What constitutes a dark pattern and a digital nudge overlaps to a large extent: Both are intended to change behavior, but one benefits the firm, while the other benefits the decision-maker (user). In competitive markets, we can conceptualize private firms' design choices as 'Pareto' or 'rent-seeking' nudges (Beggs, 2016), depending on whether the intervention improves the welfare of both the firm and the customer. Following this working definition, we turn to the experimental evidence on the effects and the prevalence of these practices. Between 15 to 20 percent of popular shopping websites contain some form of a dark pattern (Mathur et al., 2019), and with experimental work showing that common dark patterns can lead up to a fourfold increase in subscriptions to dubious services (Luguri & Strahilevitz, 2021). The most striking evidence on how easy it is to change online behavior comes from cookie consent banner designs in the wake of the General Data Protection Regulation (GDPR) in EU: Small changes in their design such as the coloring of buttons, text-size, positioning of the banner, and the order of presentation of consent options are all found impact consent rates (Graßl et al., 2021; Utz et al., 2019).

I focus on practices, identified by regulators across EU and US as causes for digital addiction.[2] The core idea I build upon is that firms optimize their platforms for consumption at the expense of their users, by strategically modifying the content served on their platforms and its design to their own liking (Ichihashi & Kim, 2022). A relevant and illustrative account –albeit anecdotal– of how this type of digital nudging worked in practice is below:

"I'm the dev that built Netflix's autoplay of the next episode. [...] When I worked there the product team at Netflix had two KPIs all new features were tested against: hours watched and retention. We would come up with all sorts of ideas to try out, and release them to small user populations of about 100,000 or so. It was great because you didn't

[2]I refer to US SMART Act of 2019, S. 2314, 116th Congress, and the upcoming EU Digital Fairness Act.

have to debate much about whether a new feature was a good idea or not, you just built it and tested it. If the feature didn't increase hours watched or retention in a statistically significant way, the feature was removed.

Autoplay massively increased hours watched. I can't remember the exact numbers, but it was by far the biggest increase in the hours watched KPI of any feature we ever tested. There was some skepticism about whether the number was inflated by Netflix continuing to play when the user left the room. [...] So yes, Netflix wants you to spend more hours watching Netflix and the product team is scientifically engineering the product to make it more addictive. But...the product team at Doritos does the same thing." (anonymous Netflix developer, 2019).

1.1.2 Methodological background

Inspired by this, my first chapter provides a laboratory experiment on the effect of a specific interface design feature – autoplay – on content consumption. I use a foundational methodological framework in economics called intertemporal choice. This framework deals with decisions whose consequences unfold over time, requiring people to weigh the costs and benefits that occur at different points in the future. Practically, intertemporal choice applies to almost all our daily decisions, from how much to save versus spend, to whether to exercise now or postpone it.

A central idea within this framework is that people often exhibit present-focused preferences (Ericson & Laibson, 2019), characterizing the tendency for people to be more impatient about rewards available right now compared to rewards available later. When faced with an immediate choice, people are more likely to select an action that generates instant satisfaction or pleasure, even if they might choose a more beneficial, but delayed, option if the decision were for a future time. Meta-analytic evidence supports the existence of present-focused preferences, with estimate sizes depending on study characteristics (Imai et al., 2021). This 'informal' description of present-focused preferences highlights that sometimes people act more impatiently for the present than they do for the future.

A classic illustration of present-focused preferences is seen in preference reversals. For instance, people might prefer to choose fruit over unhealthy snacks if that choice is for a week in the future. However, when the moment to choose arrives, the immediate urge and temptation for the snack often overrides the longer-term, healthier preference (Read & Van Leeuwen, 1998). The idea is that for future plans, people tend to select the 'healthy' or 'optimal' option, but in the immediate moment, temptation takes over and sways the decision. This immediate pull towards instant gratification or away from immediate displeasure (like postponing a tedious task) generates immediate experienced utility.

We can apply this concept to understand 'digital temptation' (Kleinberg et al., 2022), and draw an analogy between content consumption and food consumption. Just like someone might feel 'hungry' for a snack, people can be 'hungry' for content. In the long run, people would plan for an 'optimal' or 'healthy' digital consumption pattern. However, at the immediate moment of consuming content, temptation kicks in and alters these decisions. The degree to which content is tempting, and thus capable of swaying immediate choices, is shaped by what digital platforms offer and, critically, how they present it from a user interface design perspective. Focusing on the user interface design, I hypothesize that autoplay videos are more tempting than click-to-play videos, holding all else, including content, equal. If autoplay is indeed more tempting than click-to-play, people should consume more autoplay video content at the present moment than they would in the click-to-play condition.

I designed a two-day experiment to measure the temptation of autoplay and its cost to productivity, linking it to the welfare impact of dark patterns. On the first day, participants made consumption plans for the future. On the next day, they could either follow their plan or deviate from it when the time came to act. More specifically, across two days, participants determine how much time they would allocate between a productive but exhaustive task where they earn a slightly higher wage versus a leisure task that pays slightly less. At the heart of the trade-off is that labor is costlier than leisure per seconds spent at the task. My conjecture is that the combination of effortful labor and tempting leisure should lead me to observe a snack-fruit type reversal in present versus future decisions across the two dates. Crucially, I randomize my leisure offer: half the participants have autoplay and the other half have click-to-play videos, while keeping the content shown in both treatments identical.

1.1.3 Summary of findings and discussion

To my surprise, the first finding is that my decision environment cannot be characterized by present-focused preference revearsals, as most participants spend more time on the effortful labor task than they intended a day ago. Instead, my decision environment is akin to studies where participants have a preference for improving sequences (Bhatia et al., 2021; Kahneman et al., 1993; Loewenstein & Prelec, 1993) – an under examined part of the literature where people to end up pushing leisure to later consumption stages. These time preference studies found that people at times decide for having the unpleasant experience (working) first and enjoy later. Aligning with their findings, experimenter demand effects resulting from my design choices (e.g., the earnings difference between tasks) have indeed shifted attention away from the leisure task, with participants getting more work done and confounding treatment effects.

I interpret this result as a critique of my own methodological approach, which prioritized a controlled setup over creating a setting where video consumption occurs naturally. "The time preference literature is often summarized as a list of stylized facts (e.g., people are impatient, discount functions are hyperbolic, people prefer improving sequences). This characterization falsely suggests the existence of a small set of robust psychological phenomena which measurement procedures merely record. In reality, many of the widely cited, stylized 'facts' remain facts only by virtue of an unwitting convergence in research methodologies. Those studying time preferences should use more diverse measurement procedures and devote more attention to the question of how respondents resolve inconsistencies among them." (Frederick & Loewenstein, 2008, p. 232). Because the framing of questions and study design have such deep implications, a different setup could have resulted in findings that aligned with my initial hypotheses. Still, that design would have once again prioritized neatly prepared tasks –in lieu of studying the decision-making environment where actual consumption decisions occur. Frederick and Loewenstein's critique raises the question whether my time preference approach was the right way to study autoplay, and whether a simpler approach in a field experiment would have been more appropriate.

As participant attention shifted from watching leisure videos, I find that there is no difference in terms of the content consumed between the two groups who were randomized into autoplay or click-to-play conditions. An unexpected finding as evidence from the field points to an effect, where autoplay increases content consumption (Schaffner et al., 2025). I conjecture that this could point to two culprits. The first one is my methodological approach inside the lab as discussed above, which I label the context, and second, the content. Indeed, content may also play a role in what makes autoplay addictive: think animal documentaries versus celebrity gossip, where would autoplay be more 'tempting'? I deliberately served identical videos to all participants to isolate autoplay. I curated these among popular videos from YouTube and TikTok, with the topic of the videos chosen by a majority vote between pretest participants from the same population as the study sample. Despite these, my content offered no algorithmic personalization or choice. In this sense, autoplay may not be problematic if people think the video content is not tempting in the first place.

To support this conjecture, I report findings from the Willingness to Pay (WTP) block of the same experiment conducted at a later date. Participants in this second block are willing to spend a small positive amount to keep autoplay on, considering autoplay as a useful feature worth giving up earnings for. This is a second revealed preference evidence against the temptation hypothesis. Further, autoplay, by removing the explicit decision to consume the next content, speeds up consumption of content. While both conditions consume the same number of videos, participants in autoplay condition spend more time working due to this phenomenon. The amount participants are willing to pay for autoplay is very close to the amount of money lost by clicking to play each video in the absence of autoplay. These point to some efficiency gains in content consumption made by autoplay.

Combining the content with the context, I conclude that my experimental setup, where I tracked every second spent on leisure instead of a higher paid work task, ultimately failed to capture to the real-world autoplay video consumption phenomenon that I intended to study. The lessons learned from this first experiment are invaluable, and have led me to alter the type of contextual settings I wanted for the second part of my dissertation.

1.2 From the laboratory to the field

The onset of this dissertation was to examine a bundle of design features suspected to change behavior and document how they would go against users' preferences using the economic toolbox as exemplified above. Yet, the conclusion of my first chapter led me to desire a more naturalistic setting where I would be closer to the field, in the remainder of my research. Events during the second year of my program made it so that I could get experience working with experts conducting research in the field, but to do so, I needed to change my focus to studying the social and institutional constraints instead of the digital ones.

Thus, in the remainder of my dissertation, my conceptual focus switches to the social and institutional, considering a higher education setting and studying how the implicit or explicit incentive structures within it shapes students' accumulation of social capital. For both second and third chapters, I work in the same university setting in Colombia, as I have access to the student body for running experiments and admin data to supplement my analyses. For this reason, I believe the partner university and the higher education setting in Colombia merit an introduction.

Colombia has very high levels of economic inequality as measured by the Gini coefficient (United Nations, 2023; World Bank, 2024). It implies that the income difference between the richest and the poorest in the population is extreme. This translates into a segregated society where educational outcomes are very different depending on socioeconomic status (SES). In addition, the share of students enrolled in private higher education institutions is much higher than in OECD average, standing at 47% vs. 30% (Villegas, 2021). This high share implies that private schools offer a diverse range of quality, catering to both higher and lower performing students. As a consequence, children from wealthy families typically attend exclusive, high-quality private schools, while children from poorer families attend public schools or private schools of regular to poor quality.

Many families, including those from middle and lower-middle classes, choose private education even when low- or medium-cost private institutions do not offer demonstrably better academic quality than comparable public ones (Fergusson & Flórez, 2021). García Villegas and Cobo (2021) attribute this to families seeking status and social distinction through education, which provides not only cultural signals (specific ways of speaking and behaving) but also social capital by granting access to more privileged individuals.

Against this backdrop, certain non-elite but private institutions –like our partner university– cater to a large group of students, and create an opportunity to mix between the rich and poor. My research capitalizes on this diversity, where different social groups coexist within the same educational setting. I evaluate whether the opportunity to mix across SES groups translates into meaningful social outcomes in terms of referrals. The field setting therefore allows me to explore the economic consequences of private education at a key moment in student life. Going back to our initial definition, knowing high-SES (wealthy) individuals does not equal having them as contacts who would pass on opportunities, as access does not imply mobilization of social capital. The university environment simply satisfies the precondition for a relationship which may or may not evolve. Put differently, I ask whether the fees paid for private education translate into social capital.

1.3 Learning about peers when signals are mixed

In chapter two, we start by looking at how much cross-SES interaction happens within the classroom, which qualities (skills) of their peers students can observe, and whether we can remedy inequalities by redirecting referrals to low-SES classmates without compromising the performance.

1.3.1 Theoretical background and motivation

One part of our motivation comes from the advances in the human capital literature within the last two decades. Human capital theory refers to the now widely accepted idea that education, training, and other forms of learning are investments that pay off in the future (Becker, 1964/1993). To be clear, capital here consists of the resources devoted to the preparation of labor in the form of education and training. Deming (2022) summarize the recent advancements in the field under four stylized facts: First, human capital, as the sum of the learning investments made, explains a substantial share of the variation in labor earnings within and across countries. This positive relationship between schooling quantity and future earnings is one of the most accepted findings in social science. Second, human capital investments have the highest economic returns in early childhood and decrease with age, referring to the famous Heckman curve where earlier investments cascade over time (Heckman, 2006). Third, to produce more foundational skills such as numeracy and literacy, resources (money spent on education) are the main constraint. Finally, 'non-cognitive' and 'social' skills such as problem-solving, conscientiousness, and teamwork are increasingly economically valuable (Deming, 2017; Heckman & Kautz, 2012; Lindqvist & Vestman, 2011; Weinberger, 2014), but are harder to measure and develop than the foundational skills.

Relevant for us are social skills, whose returns roughly doubled in the period between 1980 and 2000 (Deming, 2017; Edin et al., 2022). In other words, workers with higher social skills –given their cognitive skill– earned increasingly more for their skills, as the share of collaborative work has increased with time. Deming (2023) identifies two issues with the measurement of social skills. First, measures of social skills consist mostly of self-report questionnaires instead of behavioral outcomes. Second, there is a problem of scope and clarity. What skills are 'social', and do all social skills lead to more returns? For example, my work on social skills in this chapter solely focuses on the ability to recognize others' emotions, a strong predictor of team performance (Weidmann & Deming, 2021). Regardless, social skills are found to be both harder to discern compared to cognitive skills by recruiters, and difficult to convey credibly by workers (Bassi & Nansamba, 2022; Caldwell & Danieli, 2018). To make things worse, people have difficulties telling lies apart from truth in face-to-face interactions (Serra-Garcia & Gneezy, 2021), which could further confound the signals from workers during the hiring process. This difficulty in observing social skills creates an information asymmetry in hiring. My research, therefore, asks whether we can solve this information asymmetry by using peer referrals to identify individuals with high social skills. To compare how well social skills are recognized in comparison to cognitive skills I will ask students to refer classmates for both.

The second part of the motivation is the contact hypothesis (Allport, 1954), which broadly describes the negative correlation between intergroup contact and prejudice (Pettigrew &

Tropp, 2006). While the original formulation specifies that reduced prejudice will result when four scope conditions for the contact situation are met,[3] recent work suggests contact itself is, in most cases, sufficient in reducing prejudice (Paluck et al., 2019). A subset of these studies leverage random assignment in college dorms, where researchers evaluate the impacts of having roommates from a different race group. Findings suggest increased support for affirmative action and frequency of contact with other members of the said group (Boisjoly et al., 2006; Carrell et al., 2019), reduced prejudice as measured by the Implicit Association Test (IAT) as well as positive impact on the academic performance of the students in the disadvantaged group (Corno et al., 2022).

By fostering connections among students from different social groups over many years, educational settings tend to create favorable, i.e., collaborative and on equal status, intergroup contact opportunities. For example, classroom exposure to poor classmates makes rich students more prosocial, and more willing to socialize with other poor students (Rao, 2019). Further, collaborative contact in the context of team sports increases cross-group friendships and time spent together, as well as reducing own-group favoritism (Lowe, 2021; Mousa, 2020). Finally, careful allocation of seats within the classroom creates more diverse friendships (Rohrer et al., 2021), and mixing tends to improve grades for low-SES students (Van Ewijk & Sleegers, 2010). These suggest intergroup contact in educational settings can attenuate cultural differences between groups, generate positive feelings of trust and empathy, and can even facilitate social mobility for the poor.

1.3.2 Methodological background

The methodological inspiration directly originates in the groundbreaking work from the literature in referral experiments. To illustrate their commonalities, I will present two of the foundational studies. In the first one, Beaman and Magruder (2012) created an objective task (with clear success and failure criteria) measuring cognitive ability and hired day laborers from the market. They first assessed the performance of the initial batch of workers using their task, looking at variables such as time to complete, successful

[3] These are equal status between groups, common goals, cooperation, and institutional support (e.g., authorities, law, norms)

completion, and incorrect attempts. At the end of the task, the workers were offered payment to return with a referral, and were specifically asked to bring in candidates who they thought would perform well at the same task. The initial batch of workers was randomized into different incentive schemes to refer, with some being offered a fixed payment regardless of their referral's performance, and others being offered a guaranteed sum plus a contingent bonus based on the referral's performance, should they bring a referral. While this setup did not allow observing the workers' entire referral choice set, i.e., all individuals among whom they can select a referral comprising their social network, we could still observe whether they decided to bring a referral, and how well their referral performed in the task. Referred workers in the experiment exhibited significantly higher performance compared to the initial batch of workers. High-performing workers were capable of selecting highly skilled referrals, but only did so when their own pay was tied to the referral's performance. The evidence from this experiment supports the idea that referrals contain productivity information when referrers are incentivized appropriately.

In the second study, Pallais and Sands (2016) used an online labor market to directly measure performance for both referred and non-referred workers. Researchers first hired a set of experienced workers to perform tasks, and solicited up to three referrals from these workers. This time, workers were not incentivized to make a referral, but had implicit incentives: making high-quality referrals could improve their relationship with the hiring firm. They then invited referred workers and a random sample of non-referred workers to apply for a different set of tasks, hiring all applicants under a given wage. The experiment randomized referrals into treatments to test whether referred workers performed better because they worked with the referrer (team production), because the reputation of the referrer was at stake (moral hazard), or simply because the referral contained information about the quality of the worker (adverse selection or learning). In the information treatment, a new firm (with a different name, location, job posting, and writing style) made direct job offers to all referred and non-referred workers from the earlier peer influence experiment sample. Workers were asked to visit the Twitter pages of three musicians and answer a 10-question survey about those accounts daily for 5 consecutive days, a task that was designed as an individual diligence task over time. Referred workers demonstrated substantially higher performance and exhibited lower turnover compared to non-referred workers. Referrals from high-performing referrers and those with stronger social ties to their referrers were found to be particularly informative. Once again, results supported the idea that referrals reveal information about the quality of the worker.

The design features common in both studies are creating an objective measure such as a test, or a job, and hiring workers from the market. Researchers then assess the performance of the first batch of workers, and then ask them –under different incentive schemes– to bring in (refer) qualified workers for the job. We modify this methodological setup and leverage our educational setting, where we collect objective performance measures from students who complete tasks to earn money. Our tasks consist of two incentivized tests for cognitive and social skill. Departing from earlier work, we first collect performance measures from all individuals who could potentially receive or send referrals, including all peers in the classroom. This way, we ensure that both classmates who make referrals and those who receive them complete the same skill tests before making referrals. We then solicit three referrals among classmates for each task, where we incentivize referrers based on their referral's performance in the skill tests. This modified setup allows us to compare those who get more referrals to the rest of the classroom.

Students in our setting, most of them who are in the beginning of their second year at the university, choose to take two compulsory attendance classes depending on the availability in their schedule. Their choices result in mixed classrooms across academic programs, and SES. After spending a semester in small classrooms with less than 30 students, we individually assess students' cognitive and social skills first and then ask for classroom referrals, paying referrers (senders) on the basis of the skill test scores of their referrals (receivers). We define by referrals a recommendation task where referrers choose the names of their referrals from a list of all classmates. This task is anonymous, and done privately: We do not contact the individuals who receive referrals. Referrers do not know how well their peers did in these tests, but we hypothesize that classroom contact results in learning about the skills of their peers. Using randomized assignment, we incentivize half of the students in the classroom to refer more among their low-SES peers, without compromising skill performance.

1.3.3 Summary of findings and discussion

We first look at whether receiving a higher share of referrals in the classroom is related to the students' skills. We find that while in class interaction results in classmates recognizing cognitive skill, social skill is not captured in referrals. One reason this could be is because of our measure of social skill, a multiple choice emotion recognition test where participants look at portrait photos of people, called the Multiracial Reading the Mind in the Eyes Test (Kim et al., 2022). It could also be that social skill is harder to observe compared to cognitive skill. Inspired by the idea that social skill is harder to observe in the classroom, we separate referrals made by the same referrer. When classmates referred the same person for both social and cognitive skills, we find that academic grades were a better predictor of receiving a referral than actual skill test results. We conjecture that in the absence of a clear signal about their peers' skills, students used grades as a proxy for both skills. For unique cognitive skill referrals (choosing the person only for a cognitive skill referral), classmates refer based on grades and cognitive skill test score alike. This suggests meaningful cognitive skill information is conveyed to classmates. For unique social skill referrals, however, we find classmates refer neither based on grades nor skill test scores, supporting the idea that classmates have no information about the social skill of their peers.

We then look at whether receiving a higher share of referrals in the classroom is dependent on student SES. We find no general SES bias in referrals for comparable students with same test scores and grades, suggesting classroom diversity is effective. There is a small SES bias in referrals in unique cognitive skill referrals, resulting in qualified low-SES getting a lower share of referrals in the classroom, but this is entirely mitigated with the additional incentives in the treatment. We support these results with an additional SES guessing task, where we find that classmates are better than chance at identifying low-SES and high-SES peers, which suggests the cross-SES interaction takes place despite students being conscious of apparent SES differences between students. These results from the point of view of the contact theory are overwhelmingly positive: When different SES groups interact inside the classroom, they tend to learn about each other's labor market relevant skills to some extent, and overall are not biased against disadvantaged low-SES peers.

To sum, this chapter is an attempt at capturing the concrete effects of intergroup contact on equal status during a reasonably long period. Our findings indicate that students know each other's academic performance and cognitive skills well enough to make successful referrals after interacting. While we do not provide a concrete before and after test for the extent of the learning, we conjecture that learning is due to time spent inside the classroom. Providing a formal test for this could be an interesting avenue for future research. Further, SES biases in the referral decisions are limited and can be remedied by easy-to-implement incentives for low-SES, which we show that do not impact referral performance in skill test scores or grades.

The crucial lesson from this chapter is that classroom interaction across SES leads to positive outcomes which are translated into, or captured by, referrals. In this sense, families paying to access private educational institutions in the Colombian setting are getting returns for their investments in the form of both accessed and mobilized social capital – so long as disadvantaged students are in the same classrooms with wealthier peers. In the next chapter, we will address the question of how often these cross-SES interactions take place at the broader –university– level, and demonstrate the extent of the social capital created when constrained by institutional factors.

1.4 Sharing opportunities via referrals and the role of institutional constraints

The questions of interest emerging from the earlier chapter revolve around the finding that classroom contact across SES has tangible benefits in terms of social capital acquisition. Yet, a single class is not sufficient to assess the extent of social capital that would be created at university. Acknowledging the need to go broader, my inquiry in this chapter encompasses students' entire university network. In chapter three, we create a university network for the students at the partner institution, and combine it with a network-wide referral task instead. Methodologically, our design closely resembles the earlier experiment in chapter two.

1.4.1 Theoretical background and motivation

The motivation for this study is grounded in the literature on homophily in social networks, defined as the tendency of similar individuals across observable characteristics –including SES– to connect more often, with the intensity of association getting stronger as the similarity between individuals increases (McPherson et al., 2001). Homophily implies higher levels of similarity between closer ties in their frequency of interaction, shared sentiment, and shared resources; with individuals who are close in relations having a tendency for similarity also in terms of social capital (Lin, 2008). When individuals seek returns in resources, as in wealth, power, or a job, the 'value' of their network is conditional on the quality of its embedded resources. If the individual is relatively poor in resources, the homophily principle dictates that their closer ties will also likely involve connections with relatively poor resources. In this sense, homophily may lead to networks of SES groups within the same university to segregate and accumulate social capital differently, despite classroom interactions when taking the same classes. It would have consequences for 'accessing' and 'mobilizing' social capital in university networks.

When 'mobilizing' social capital by making referrals from the university networks, homophily can materialize in two ways (Currarini et al., 2010). First, homophily can manifest itself as a bias in the chances of referring a different 'type' of individual (low- or high-SES), equivalent to picking an individual at random from either group. We can measure how dissimilar the SES compositions of low- and high-SES networks are: the group-shares of each type in the networks of low- and high-SES determines the extent of this bias. This type of homophily closely follows the 'access' principle, where not meeting members of the other group lowers the chances to refer them. Second, given the group-shares in the university networks, homophily could manifest itself as a bias in the preferences of individuals over the types of their referrals. This is dependent on whom a student refers from their university network. At the group level, we can measure how dissimilar actualized referrals are in comparison to the average university network compositions, and determine the extent of this bias. Our goal is to characterize the extent and the role of both types of SES-homophily when 'mobilizing' social capital at university.

1.4.2 Methodological background

The methodological foundation for creating university networks follows Kossinets and Watts (2009), who build student networks based on the shared classes between students. A connection between two students is created if they took the same class during one semester, and the number of shared classes students take together define the tie strength. Following our earlier definition of social capital along with Chetty and colleagues' work, we consider how large an individual's university network is, what share of their network is composed of high-SES peers, and how strong the relationship between the network members are. To go beyond descriptive characterization of university networks, we conduct a referral task within these. This time, we ask for referrals based on the nationwide university entry exam. We have entry exam scores and SES data for all individuals who send referrals (all participants who join the experiment) and those who could potentially receive them (all peers within the participant's university network). From this rich dataset, we compare the outcomes of those who get referrals with the rest of the university network.

Like our earlier experiment in chapter two, we asked for referrals with an objective performance measure, with incentives for the referrers (sender) depending on their referrals' (receiver) score in the university entry exam. We randomized half of the referrers to a condition where their referral would earn a sizable monetary bonus. The treatment was set to evaluate how the additional bonus would impact referral performance and tie strength between the referrers and their referrals.

1.4.3 Summary of findings and discussion

We first understand how referrals are made from the university networks. Despite having a large number of potential referral candidates, in contrast to the classroom setting of chapter two, we find that referrals consistently go to higher performing students with whom the referrer has taken a very high number of classes together. Surprisingly, the additional bonus for the referrals has no effect on neither the referral's performance in the exam nor the number of classes taken together with the referrer. We interpret this as referrals going to close social contacts regardless of the incentive structure. Aggregating referrals made by SES groups, we find a significant same-SES bias for low-SES referrers but not for other SES groups. In other words, we only observe bias in the preferences of low-SES over the types of their referrals. This is a positive note on how referrals are not putting low-SES at a disadvantage even on the larger, university scale.

Yet, the lack of bias in the preferences does not imply that referrals flow from high-SES to low-SES in abundance and that the intergroup diversity works flawlessly for increasing social capital. With SES groups having similar shares in networks compared to the university averages, network compositions are overall balanced. However, this balance disappears when looking at stronger ties. As the number of shared classes between students increases, the share of same-SES peers in their networks dramatically rises. In other words, the networks of different SES groups segregate as students' relationships strengthen through repeated interaction. In the range of tie strength where most referrals occur, high-SES students have about half as many low-SES peers in their networks compared to the university average. In terms of Currarini and colleagues' language, the chances to meet at this part of the network is so low that referrals to low-SES are almost non-existent, even without the bias in referral preferences.

What causes these dramatic changes in network compositions? With 93% of referrals going to peers within the same academic program, and the number of electives within programs being at most 25% of all courses, program selection single-handedly drives the differences in network compositions as the number of courses taken together increases. The university prices programs based on their cost –with differences in yearly fees reaching up to 6 times between some programs. Further, less than 5% of students receive any scholarships. Looking at the distribution of students at the university across these programs, we find strong evidence for SES-based program selection: Low-SES sort into affordable programs, and high-SES are dominating the more expensive ones.

Institutional constraints, in the form of program fees, play a crucial role in segregating students at university by SES. The combination of fixed-costs (program-fees) and financial pressure results in suboptimal levels of investment for low-SES, impacting even occupational choices (Galor, 2011). Becker (1964/1993) illustrates the argument in four points. First, it is inherently difficult to borrow funds to invest in human capital because it cannot be offered as collateral. Second, investment often depends on internal financing (such as family resources). Third, for individuals without sufficient family wealth, the necessity of external funding to cover tuition fees and foregone earnings (earnings difference between earlier vs. later entry to the labor market) drives up their costs. Finally, poorer families may even need to finance investments by reducing their own consumption, which further discourages expenditure on children. Therefore, when investments are constrained by limited access to funds, poorer families are prevented from making the wealth-maximizing investment in their children's human capital. This situation leads to lower total investment in education for low-SES individuals and implies that the marginal rate of return on their investments tend to be higher than in richer families. To sum, the differences in the availability and cost of funds not only limits the accumulation of human capital. As we have seen it also has implications for social capital.

Based on the literature, I briefly discuss two policy solutions for future research. One solution typically risks adding burden to the public sector or reducing profit margins of the private institutions as it requires financing. It consists of providing conditional cash transfers (CCTs) or educational vouchers to disadvantaged students. These are financial support schemes to low-SES households on the condition that they invest in the human capital of children. Meta-analytic evidence consistently shows positive effects of CCTs on enrollment levels and attendance, while evidence on learning outcomes, test scores, and future earnings remains mixed (Garcia & Saavedra, 2023). Similarly, results from educational voucher interventions are inconclusive, with effects varying significantly across institutional contexts and implementations (E. Bettinger, 2011). The substantial costs and mixed results on educational performance stemming from these programs have prompted researchers and policymakers to explore different options.

The university admissions and financial aid application processes represent a significant administrative burden which can be especially challenging to navigate for disadvantaged students. Further, various financial support schemes are already in place in many higher education institutions, including in our partner university. The second solution centers on overcoming institutional barriers with low-cost behavioral interventions, focusing on simplifying communication, correcting beliefs, and setting effective defaults. Consisting of non-financial interventions, these focus on better allocating existing resources to eligible candidates. The current consensus among researchers is that they are promising, National Academies of Sciences, Engineering, and Medicine (2023) and Dynarski et al. (2023) both reporting robust successes across various outcomes. Most relevant to our context are the increases in student aid take-up by information provision and assistance (E. P. Bettinger et al., 2012; Herber, 2018; Hoxby & Turner, 2013), as well as correcting misperceptions about eligibility status (Riedmiller, 2025). Because these interventions provide easy to export and low-cost solutions to increase take-up in existing student aid programs, I recommend future research to try implementing these at first in order to address inequalities in higher education, both in terms of human capital and social capital acquisition.

To sum, this chapter attempts to capture the aggregate effects of intergroup contact across SES at university. Our findings indicate that students from different SES groups have limited interactions throughout their university experience. While the lack of bias in individual preferences against low-SES referral is indeed a positive outcome, the cascading institutional structures that limit interaction between low- and high-SES prevent the creation of meaningful connections between the two groups that result in the transfer of beneficial opportunities. In Bourdieu's definition, these valuable high-SES connections that the university environment provides cannot be leveraged by the low-SES, and thus, gains in social capital from the university are limited.

The main takeaway from the last two chapters is that while classroom interaction across SES leads to positive outcomes, the opportunities for different SES individuals to interact at the university level are severely limited. At least in the ways we have measured social capital, families paying to access private educational institutions in the Colombian setting may get much higher returns for their investments if institutional changes can be made to increase the frequency of classroom interactions with wealthier peers. Future research should address how cross-SES interactions can take place more frequently at the university level, as the evidence collected in these chapters demonstrate the benefits in terms of social capital for disadvantaged individuals.

The remainder of the dissertation consist of individual chapters following the order in which they were discussed here. I begin with the chapter on 'temptation and digital nudges', continue with the chapter on 'learning about peers when signals are mixed', and close with the chapter on 'sharing opportunities via referrals and the role of institutional constraints'. I then provide a brief conclusion for the entire dissertation.

References

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception.* Princeton: Princeton University Press.

Allport, G. W. (1954). *The nature of prejudice.* Cambridge, MA: Addison-Wesley.

anonymous Netflix developer. (2019, July 30). *Personal account of building netflix's autoplay feature.* Hacker News comment. Retrieved from https://news.ycombinator.com/item?id=20566514 (Comment by user rsweeney21 in discussion thread "New bill would ban autoplay videos and endless scrolling")

Bassi, V., & Nansamba, A. (2022). Screening and signalling non-cognitive skills: experimental evidence from uganda. *The Economic Journal*, *132*(642), 471–511.

Beaman, L. (2016). Social Networks and the Labor Market. In *The Oxford Handbook of the Economics of Networks* (pp. 648–672). Oxford University Press. doi: 10.1093/oxfordhb/9780199948277.013.30

Beaman, L., & Magruder, J. (2012, December). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, *102*(7), 3574–3593. doi: 10.1257/aer.102.7.3574

Becker, G. S. (1964/1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). Chicago: The University of Chicago Press.

Beggs, J. N. (2016). Private-Sector Nudging: The Good, the Bad, and the Uncertain. In S. Abdukadirov (Ed.), *Nudge Theory in Action* (pp. 125–158). Cham: Springer International Publishing. doi: 10.1007/978-3-319-31319-1_6

Bettinger, E. (2011). Educational vouchers in international contexts. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 551–572). Amsterdam: Elsevier. doi: 10.1016/B978-0-444-53444-6.00007-9

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R block FAFSA experiment. *Quarterly Journal of Economics*, *127*(3), 1205–1242. doi: 10.1093/qje/qjs017

Bhatia, S., Crawford, M. M., McDonald, R. L., Moreno, M. A., & Read, D. (2021, May). *Inconsistent Planning and the Allocation of Tasks Over Time* (OSF Preprints No. b4mg7). Center for Open Science.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., & Eccles, J. (2006, November).

Empathy or Antipathy? The Impact of Diversity. *American Economic Review*, *96*(5), 1890–1905. doi: 10.1257/aer.96.5.1890

Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfattheicher, S. (2016, October). Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. *Proceedings on Privacy Enhancing Technologies*, *2016*(4), 237–254. doi: 10.1515/popets-2016-0038

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood Press.

Caldwell, S., & Danieli, A. (2018). Overcoming information asymmetry in job search: The power of a reference letter. *Working Paper*.

Carrell, S. E., Hoekstra, M., & West, J. E. (2019, November). The Impact of College Diversity on Behavior toward Minorities. *American Economic Journal: Economic Policy*, *11*(4), 159–182. doi: 10.1257/pol.20170069

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... Wernerfelt, N. (2022). Social capital 1: Measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121. doi: 10.1038/s41586-022-04996-4

Corno, L., La Ferrara, E., & Burns, J. (2022, December). Interaction, Stereotypes, and Performance: Evidence from South Africa. *American Economic Review*, *112*(12), 3848–3875. doi: 10.1257/aer.20181805

Currarini, S., Jackson, M. O., & Pin, P. (2010, March). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National Academy of Sciences*, *107*(11), 4857–4861. doi: 10.1073/pnas.0911793107

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The quarterly journal of economics*, *132*(4), 1593–1640.

Deming, D. J. (2022). Four Facts about Human Capital. *Journal of Economic Perspectives*, *36*(3), 75–102. doi: 10.1257/jep.36.3.75

Deming, D. J. (2023). Chapter 6 - Multidimensional human capital and the wage structure. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 7, pp. 469–504). Elsevier. doi: 10.1016/bs.hesedu.2023.03.005

Dynarski, S., Nurshatayeva, A., Page, L. C., & Scott-Clayton, J. (2023). Addressing nonfinancial barriers to college access and success: Evidence and policy implications. In E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 6, pp. 319–403). Amsterdam: Elsevier.

Edin, P.-A., Fredriksson, P., Nybom, M., & Öckert, B. (2022). The Rising Return to Noncognitive Skill. *American Economic Journal: Applied Economics*, *14*(2), 78–100. doi: 10.1257/app.20190199

Ericson, K. M., & Laibson, D. (2019). Intertemporal choice. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 1–67). Elsevier. doi: 10.1016/bs.hesbe.2018.12.001

Fergusson, L., & Flórez, S. A. (2021). Desigualdad educativa en colombia. In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel.

Frederick, S., & Loewenstein, G. (2008, December). Conflicting motives in evaluations of sequences. *Journal of Risk and Uncertainty*, *37*(2-3), 221–235. doi: 10.1007/s11166-008-9051-z

Galor, O. (2011). Inequality, human capital formation, and the process of development. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4, pp. 441–493). Amsterdam: Elsevier.

Garcia, S., & Saavedra, J. E. (2023). Conditional cash transfers for education. In E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 6, pp. 499–590). Amsterdam: Elsevier. Retrieved from https://doi.org/10.1016/bs.hesedu.2022.11.004 doi: 10.1016/bs.hesedu.2022.11.004

García Villegas, M., & Cobo, P. (2021). La dimensión cultural del apartheid educativo. In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel.

Graßl, P., Schraffenberger, H., Borgesius, F. Z., & Buijzen, M. (2021). Dark and Bright Patterns in Cookie Consent Requests. , 39.

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018, April). The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). Montreal QC Canada: ACM.

doi: 10.1145/3173574.3174108

Hanson, J. D., & Kysar, D. A. (1999). Taking Behavioralism Seriously: The Problem of Market Manipulation. *New York University Law Review*, *74*, 630.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, *312*(5782), 1900–1902. doi: 10.1126/science.1128898

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464. doi: 10.1016/j.labeco.2012.05.014

Herber, S. P. (2018, February). The role of information in the application for highly selective scholarships: Evidence from a randomized field experiment. *Economics of Education Review*, *62*, 287–301. doi: 10.1016/j.econedurev.2017.12.001

Hoxby, C. M., & Turner, S. (2013). *Expanding college opportunities for high-achieving, low income students* (SIEPR Discussion Paper No. 12-014). Stanford Institute for Economic Policy Research.

Ichihashi, S., & Kim, B.-C. (2022). Addictive Platforms. *Management Science*. doi: 10.1287/mnsc.2022.4392

Imai, T., Rutter, T. A., & Camerer, C. F. (2021, May). Meta-Analysis of Present-Bias Estimation using Convex Time Budgets. *The Economic Journal*, *131*(636), 1788–1814. doi: 10.1093/ej/ueaa115

Jackson, M. O. (2008). Social and Economic Networks.

Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., ... Weber, E. U. (2012, June). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*(2), 487–504. doi: 10.1007/s11002-012-9186-1

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993, November). When More Pain Is Preferred to Less: Adding a Better End. *Psychological Science*, *4*(6), 401–405. doi: 10.1111/j.1467-9280.1993.tb00589.x

Kim, H., Kaduthodil, J., Strong, R. W., Germine, L., Cohan, S., & Wilmer, J. B. (2022, August). *Multiracial Reading the Mind in the Eyes Test (MRMET): An inclusive version of an influential measure* (Preprint). Open Science Framework. doi: 10.31219/osf.io/y8djm

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2022). *The Challenge of Understanding What Users Want: Inconsistent Preferences and Engagement Optimization* (No. arXiv:2202.11776). arXiv.

Kossinets, G., & Watts, D. J. (2009, September). Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*, *115*(2), 405–450. doi: 10.1086/599247

Lin, N. (2008). A network theory of social capital. In D. Castiglione, J. W. van Deth, & G. Wolleb (Eds.), *The handbook of social capital* (pp. 50–69). Oxford: Oxford University Press.

Lindqvist, E., & Vestman, R. (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–128. doi: 10.1257/app.3.1.101

Loewenstein, G. F., & Prelec, D. (1993, January). Preferences for sequences of outcomes. *Psychological Review*, *100*(1), 91–108. doi: 10.1037/0033-295X.100.1.91

Lowe, M. (2021, June). Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration. *American Economic Review*, *111*(6), 1807–1844. doi: 10.1257/aer.20191780

Luguri, J., & Strahilevitz, L. J. (2021). Shining a Light on Dark Patterns. *Journal of Legal Analysis*, *13*(1), 43–109. doi: 10.1093/jla/laaa006

Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019, November). Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–32. doi: 10.1145/3359183

Mathur, A., Kshirsagar, M., & Mayer, J. (2021, May). What Makes a Dark Pattern... Dark?: Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). Yokohama Japan: ACM. doi: 10.1145/3411764.3445610

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*, 415–444.

Mousa, S. (2020, August). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science*, *369*(6505), 866–870. doi: 10.1126/science.abb3153

National Academies of Sciences, Engineering, and Medicine. (2023). *Behavioral economics: Policy impact and future directions.* Washington, DC: The National Academies Press. Retrieved from https://doi.org/10.17226/26874 doi:

10.17226/26874

Pallais, A., & Sands, E. G. (2016, December). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, *124*(6), 1793–1828. doi: 10.1086/688850

Paluck, E. L., Green, S. A., & Green, D. P. (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy*, *3*(2), 129–158. doi: 10.1017/bpp.2018.25

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783.

Rao, G. (2019, March). Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools. *American Economic Review*, *109*(3), 774–809. doi: 10.1257/aer.20180044

Read, D., & Van Leeuwen, B. (1998, November). Predicting Hunger: The Effects of Appetite and Delay on Choice. *Organizational Behavior and Human Decision Processes*, *76*(2), 189–205. doi: 10.1006/obhd.1998.2803

Riedmiller, S. (2025). *Reducing Inequality by Correcting Misperceptions: Experimental Evidence on Student Aid Take-Up.* SSRN. doi: 10.2139/ssrn.5364685

Rohrer, J. M., Keller, T., & Elwert, F. (2021, August). Proximity can induce diverse friendships: A large randomized classroom experiment. *PLOS ONE*, *16*(8), e0255097. doi: 10.1371/journal.pone.0255097

Schaffner, B., Ulloa, Y., Sahni, R., Li, J., Cohen, A. K., Messier, N., … Chetty, M. (2025, May). An Experimental Study Of Netflix Use and the Effects of Autoplay on Watching Behaviors. *Proc. ACM Hum.-Comput. Interact.*, *9*(2), CSCW030:1–CSCW030:22. doi: 10.1145/3710928

Serra-Garcia, M., & Gneezy, U. (2021, October). Mistakes, Overconfidence, and the Effect of Sharing on Detecting Lies. *American Economic Review*, *111*(10), 3160–3183. doi: 10.1257/aer.20191295

Thaler, R. H. (2017). Behavioral Economics. *Journal of Political Economy*, *125*(6), 1799–1805. doi: 10.1086/694640

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness.* Penguin.

Topa, G. (2019, December). Social and spatial networks in labour markets. *Oxford Review of Economic Policy*, *35*(4), 722–745. doi: 10.1093/oxrep/grz019

United Nations. (2023). *Social panorama of latin america and the caribbean 2023: labour inclusion as a key axis of inclusive social development.* ECLAC and United Nations. Retrieved from https://www.cepal.org/es/publicaciones/68702-panorama-social-america-latina-caribe-2023-la-inclusion-laboral-como-eje-central

Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019, November). (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (pp. 973–990). London United Kingdom: ACM. doi: 10.1145/3319535.3354212

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134–150. doi: 10.1016/j.edurev.2010.02.001

Villegas, M. G. (2021). *La quinta puerta.* Ariel Colombia.

Weidmann, B., & Deming, D. J. (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica*, *89*(6), 2637–2657. doi: 10.3982/ECTA18461

Weinberger, C. J. (2014). The Increasing Complementarity between Cognitive and Social Skills. *The Review of Economics and Statistics*, *96*(5), 849–861. doi: 10.1162/REST_a_00449

Wendel, S. (2016). Behavioral nudges and consumer technology. In S. Abdukadirov (Ed.), *Nudge theory in action: Behavioral design in policy and markets* (chap. 5). Cham: Palgrave Macmillan.

World Bank. (2024). *Regional poverty and inequality update spring 2024* (Poverty and Equity Global Practice Brief). Washington, D.C.: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269

Chapter 2

Does Autoplay Drive Excessive Screen Time? Evidence from an Online Experiment

Status: Not sent out for publication

Contribution Statement: Reha Tuncer:[1] Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization.

Abstract

Interface design features can 'nudge' consumers to take certain actions and are often accused of promoting addictive online behavior. A prevalent design feature across popular social media and streaming platforms is the autoplay default. In this study, I present an incentivized online experiment investigating whether the autoplay feature can cause an increase in undesired video consumption, and elicit the willingness to pay for commitment against autoplay. In a two-day study, I recruited a total of 236 participants to allocate 20 minutes between two tasks: Transcribing meaningless characters and watching funny animal videos. Time allocation decisions were planned a day before and realized on the next day. I randomly assigned participants to either autoplay or click-to-play media controls while keeping the video content constant. I find that the autoplay feature, in isolation, does not override participants' planned time allocation for media consumption. In addition, participants exhibit a positive willingness-to-pay for autoplay (6.72 pence/hour), perceiving it as a convenience feature rather than a self-control problem. Experimenter demand effects and lack of content appeal result in participants allocating more time to the transcription task than planned, confounding the effect of the autoplay treatment. These results suggest that design features promoting potentially addictive behavior like autoplay are better studied in field settings where content consumption occurs naturally alongside algorithmic personalization.

[1] I am grateful to Ernesto Reuben, Kerstin Bongard-Blanchy, and Vincent Koenig for their guidance in the conception of this experiment, to Sophie Doublet for her help in interface design, and to Suvadeep Mukherjee for his guidance in building web experiments. I also thank participants at the Behavioral Economics Summer School at the University of Copenhagen, the doctoral workshop at UC Louvain, the Experimetrics Workshop Soleto at Sapienza University Rome, as well as Samuel Greiff, Manuel Munoz, Hande Erkut and numerous colleagues for helpful comments. This study is supported by the Luxembourg National Research Fund (FNR) PRIDE 19/14302992. All errors remain my own.

JEL Classification: C91, D12, L82, D90

Keywords: experimental economics, digital addiction, interface design, self-control, media consumption

2.1 Introduction

Autoplay is a design feature that automatically provides users with new video content without any action. It has become a familiar sight in the user interfaces of major social media and streaming platforms over the last decade,[2] and has been described as a default nudge that reduces the autonomy of users (Lukoff et al., 2021; Schaffner et al., 2023). Autoplay's widespread adoption and concerns about its restriction of users' freedom of choice have also attracted regulatory scrutiny, with policymakers from both sides of the Atlantic proposing to ban platforms from serving autoplay content due to its contribution to addictive online behavior.[3]

While autoplay is pervasive and tends to increase viewing times (Chen et al., 2025; Hiniker et al., 2018; Schaffner et al., 2025), its impact on user welfare and the mechanisms driving this effect are not well understood. Does the increase in consumption go against users' wishes, or does it merely remove frictions to help them achieve their desired consumption levels? Can we attribute the increase in consumption to autoplay in isolation, or does it also depend on the supply of content? To isolate the causal effect of autoplay on content consumption and its welfare consequences, I designed an experiment holding constant the video content displayed to users while eliciting their preferences related to content consumption before the actual consumption occurs.

I dedicated the first block of the experiment to measure the differences between planned and realized allocations. I recruited 184 participants who allocated 1200 seconds (20 minutes) between two tasks. The tasks involved seconds of transcription of random

[2]As of 2025 these platforms include Facebook, Instagram, Netflix, YouTube, TikTok, and Twitter (currently X). Twitter first introduced autoplay in 2015, and described it as a means to reduce "extra effort" in the number of clicks/taps required to consume content.

[3]See US SMART Act of 2019, S. 2314, 116th Congress, and the upcoming EU Digital Fairness Act currently in discussion.

character sequences at 0.15 pence/second, and seconds of watching funny animal videos at 0.1 pence/second. On the first day, participants got familiar with tasks and made planned allocations using the strategy method. On the following day, participants realized their time allocation decision without constraints.

To isolate the causal effect of autoplay on the realized time allocation decisions, I randomly assigned participants to either the *Autoplay* or *Control* condition. In the *Autoplay* condition watching task videos played continuously. In the *Control* condition participants explicitly clicked to play each video. I held constant the content by presenting 80 manually curated "funny animal" videos from YouTube and TikTok in identical sequence. This eliminated algorithmic content curation and personalization as potential confounds.

In a second block of the experiment, I modified the design to elicit demand for a commitment device. I recruited 52 additional participants who made binary choices over their preferred autoplay setting (on or off) for nine bonus payments ranging from 5 to 50 pence. I implemented one of these decisions at random, determining both the autoplay setting for the next day and the associated bonus payment. This allowed me to measure participants' Willingness To Pay (WTP) for autoplay in advance, providing insight into whether participants viewed the feature as beneficial (positive WTP) or as a self-control problem requiring a commitment device (negative WTP).

I report five key findings. The first set of results focuses on my preregistered[4] hypothesis that autoplay would cause an increase in video content consumption. I find no evidence supporting this hypothesis. Participants in the *Autoplay* condition watched statistically indistinguishable numbers of videos compared to the *Control* condition. *Autoplay* condition also had no significant effect on the proportion of time spent on either task. These results also hold for other engagement metrics (i.e., average transcribing/watching session length, or the number of transcribing/watching sessions).

Second, participants viewed autoplay as beneficial rather than problematic. Contrary to my hypothesis that participants would demand commitment devices against autoplay, I find a positive WTP for the feature. Using data from the second block, I estimate an average WTP of 6.72 pence/hour. Autoplay is perceived as a convenience feature rather

[4] Access the preregistered hypotheses.

than as a threat to self-control, aligning with earlier research (Bongard-Blanchy et al., 2021; Schaffner et al., 2025).

Third, comparing planned time allocations to realized ones, I find that participants consumed significantly less content than their day 1 plans implied in both conditions. The decrease amounts to 35% less time spent on the watching task. This result is driven by 33% of all participants who had planned to spend up to three quarters of their time watching videos but instead decreased their content consumption on the next day. On the other end, 88% of those who had initially planned to spend less than a quarter of their time watching videos successfully followed their plans, matching their planned time allocations.

Fourth, evidence from a regression discontinuity analysis shows that participants monitored their day 1 plans despite ultimately exceeding them. When participants reached their planned transcribing duration, they immediately reduced transcribing probability by 25.5 percentage points in the next 60-second window. However, this effect dissipated quickly and became statistically insignificant within 90 seconds as participants resumed transcribing.

Finally, a mixed-methods analysis of open-ended responses reveals that participants perceived the experimental environment as a work setting rather than a choice between independent tasks. While 29% spontaneously described transcribing using negative terms, another 23% characterized video-watching as "taking breaks" from work. This suggests that experimenter demand effects may have confounded the intended effects of the autoplay treatment, with implications for both the interpretation of our null results and the broader generalizability of our study.

A primary internal validity concern is that experimenter demand effects favored allocating more time to the transcribing task than planned. Such effects can emerge through contextual cues and alter participant behavior by drawing attention to variables of interest (Zizzo, 2010). In my experiment, these could include (i) the small earnings difference favoring the transcribing task,[5] (ii) multiple mentions of the quality criteria for the tran-

[5]This difference is 60 pence over 20 minutes, equivalent to a 1.8 pound hourly wage difference (15% above our 12 pound minimum wage baseline). I hypothesized that in the absence of an earnings difference,

scribing task to receive earnings, (iii) placement of the transcribing task as the first activity upon starting a session, and (iv) the time allocation slider framing decisions as time spent on the transcription task. Alternative designs could have addressed these concerns by rewarding balanced time allocations more, eliminating earning differences altogether, or displaying both tasks right next to each other.

Another major challenge in isolating the effects of autoplay is the role of content appeal. Higher content appeal would make the watching task more tempting, and create a stronger tradeoff between the two tasks. To address this while keeping the content identical across conditions, I conducted pretests from the same participant pool where the funny animal videos came out as the most popular theme. Participants also rated the curated videos positively (median 7/10) in a post experiment survey. Regardless, the watching task content may not have been sufficiently appealing. Allowing participants to select among a list of themes with manually curated videos could have addressed this issue (Ek & Samahita, 2023).

A final concern in online experiments is ensuring genuine participation equivalent to laboratory conditions. I implemented several measures to address this challenge. I used JavaScript-based attention monitoring to detect each second participants navigated away from the experimental interface following (Purohit et al., 2023). I also required minimum internet speeds and restricted participation to desktop users with Chromium-based browsers to standardize the user experience. Participants also answered post-experiment survey questions about their engagement while watching videos and about connection interruptions. These controls ensured that my final dataset is composed of participants who were active, and the online environment could approximate the conditions of a laboratory experiment (Arechar et al., 2018).

This study contributes to several strands of literature. First, the literature studying the effects of social media and streaming services on user well-being (Allcott et al., 2020, 2022; Bao, 2025; Beknazar-Yuzbashev et al., 2025; Braghieri et al., 2022; Bursztyn et al., 2023; Collis & Eggers, 2022; Groshek et al., 2018; Nyhan et al., 2023; Purohit et al., 2023; Walton-Pattison et al., 2018). I contribute to this work by shedding light on one potential

participants would not work on the transcription task.

driver of the harmful effects of social media and streaming services, the autoplay, in a controlled setting.

A subset of this literature focuses on the effects of specific interface design elements on user behavior, particularly related to user agency and overconsumption of content (Bongard-Blanchy et al., 2021; Chen et al., 2025; Hoong, 2021; Lukoff et al., 2021; Lupiáñez-Villanueva et al., 2022; Lyngs et al., 2019, 2020; Mathur et al., 2021; Schaffner et al., 2023, 2025; Silverman et al., 2024). To my knowledge, this paper provides the first attempt at isolating the causal effects of autoplay on content consumption.

Methodologically, this paper contributes to the literature on laboratory experiments that observe behavior in real time while participants trade cognitive effort with leisure (Bhatia et al., 2021; Bonein & Denant-Boèmont, 2015; Ek & Samahita, 2023; Houser et al., 2018; Kool & Botvinick, 2014). It also adds to a burgeoning literature manipulating design elements within the user interface to reduce overconsumption of content (Hiniker et al., 2018; Lyngs et al., 2020; Purohit et al., 2023; Schaffner et al., 2025).

The remainder of this paper is organized as follows. Section 4.4 describes the experimental design. Section 2.3 details the experimental sample and procedures. Section 4.6 presents results and Section 4.8 concludes.

2.2 Design

2.2.1 Timeline and Session Structure

I conducted a two-day online lab experiment to examine how autoplay feature affects time allocation between transcription of random characters and watching funny animal videos. The multi-day structure allowed us to measure the gap between planned and actual behavior, a key feature in research on present-focused preferences (Ericson & Laibson, 2019). Participants completed two separate data collection sessions with a cooling-off period of 24 hours between sessions.

Day 1 sessions included a practice session to familiarize participants with the experimental interface, followed by planned time allocation decisions. Day 2 featured the main session where participants made actual time allocation decisions. I conducted two waves of data collection during May 2023, as described in Table 2.1.

Table 2.1: Summary of timeline and session structure

	Practice Session	Multiple Price List	Time Choice	Main Session	Decision Horizon
Day 1 (Block 1)	✓		✓		24-hour
Day 2 (Block 1)				✓	immediate
Day 1 (Block 2)	✓	✓	✓		24-hour
Day 2 (Block 2)				✓	immediate

Note: This table shows the experimental components administered across two data collection waves, with columns ordered chronologically within each day. The first block established differences in planned and actual time allocations across conditions, while the second block incorporated the Multiple Price List to measure demand for commitment against autoplay. Decision horizon refers to the period between planned time allocation and the main session.

2.2.2 Technical Implementation and Interface Design

I designed the experimental platform to provide precise control over the testing environment. I presented both tasks within a single web page using distinct tabs, allowing participants to switch between transcribing and watching videos seamlessly while preserving progress. A progress bar and timer provided feedback for remaining time. Sessions began with the transcription task as the default, requiring an active choice to switch to the watching task.

The interface incorporated several key features to track behavior. I logged all critical interactions including tab switches, keystrokes during transcription, video controls, and cursor movements outside the experimental window. This granular tracking enabled me to construct detailed measures of attention and engagement across tasks.

2.2.3 Internal Validity Controls

Online experiments present unique challenges for maintaining experimental control, including participant multitasking, technical disruptions, and environmental variation. I implemented comprehensive measures to address these concerns and ensure data quality. I clearly communicated these restrictions to participants at the experiment's outset.

Participant and Technology Screening: I selectively admitted participants using multiple screening criteria to ensure a stable experience. Internet speed requirements filtered out participants with unstable connections (minimum 30 Mbps for block 1, 40 for block 2). I restricted participation to users with Windows, Linux, or Mac operating systems using Chromium-based browsers exclusively. I excluded Safari and Mozilla Firefox due to restrictive autoplay configurations. Mobile devices and tablets were prohibited to standardize the user experience.

Platform Controls: The experimental website enforced linear navigation, preventing participants from moving backward through the experiment. I disabled right-click functions and keyboard shortcuts. Any attempt to navigate backward reset progress to the landing page.

Attention and Engagement Monitoring: I implemented real-time tracking of participant attention using a JavaScript library to detect interface visibility disruptions.[6] This system captured actions such as minimizing the browser, switching between web pages, or launching other applications. The tracking provided second-by-second measures of participant engagement.

Post-Experiment Validation: After completing the experiment, participants reported any connectivity disruptions and confirmed their genuine engagement with the watching task. I combined these self-reports with tracking measures to identify and filter data potentially compromised by distractions or inattention.

[6] See more information about the Intersection Observer API.

2.2.4 Tasks

Transcription Task

The transcription task requires transcribing successive CAPTCHAs. CAPTCHAs are computer-generated images that contain random letters, numbers and randomly placed white spaces.[7] Each CAPTCHA has a total length of 35 characters, including the white spaces. Participants spent on average 41 seconds per CAPTCHA.

I designed the images to be blurred and distorted by lines and dots added on top to increase comprehension difficulty.[8] The interface always presents these CAPTCHA images with a text box and a submit button. When the submit button is clicked, the next CAPTCHA image appears and the text input for the previous one is stored. The task is identical across conditions, as illustrated in Figure 2.1.

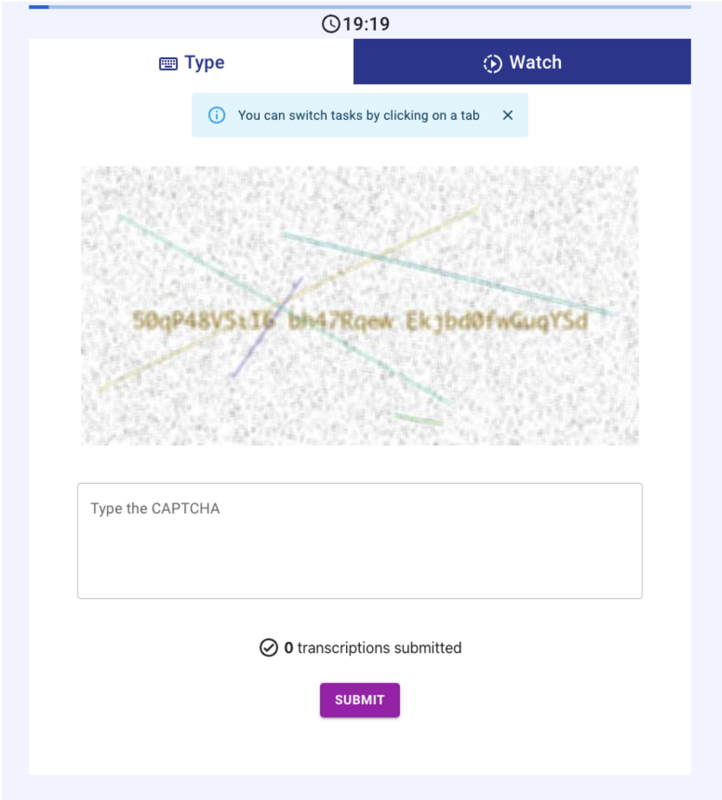
I established easy-to-meet quality requirements to ensure participants remained active. These requirements include an overall transcription accuracy of at least 70 percent in submitted CAPTCHAs[9] and a minimum of one CAPTCHA submission per minute spent on the task. Participants had to meet these quality requirements to receive compensation for the transcription task.

[7] I added white spaces to improve readability after pretests. The number of white spaces was limited to a maximum of 5 instances per CAPTCHA to control the transcribing time per image.

[8] The randomization procedure to generate CAPTCHAs.

[9] I use Python's `difflib` string-matching algorithm to measure accuracy. It consists of finding the longest common substring and then finding recursively the number of matching characters in the non-matching regions on both sides of the longest common substring. Overall accuracy of 70 percent implies an average of 10.5 $(0.3 * 35)$ mistakes across all submitted CAPTCHAs.

Figure 2.1: Transcription task tab



Note: This figure describes the transcription interface tab. As participants always begin in the transcription tab by default, there is a banner that reminds participants to click the tab buttons to switch tasks. This banner disappeared as soon as participants switched tasks once.

Watching Task

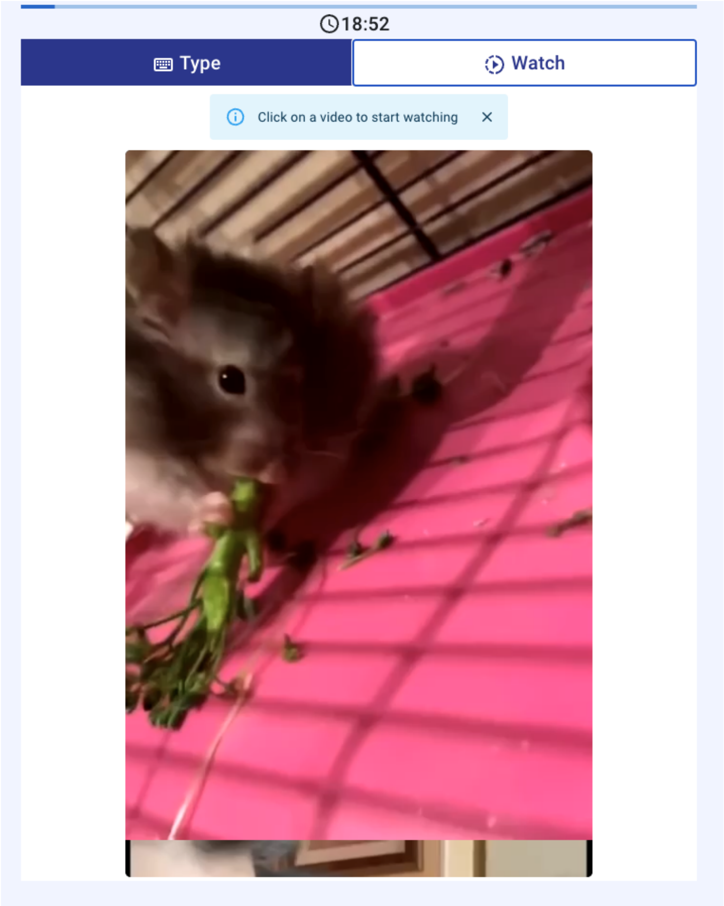
In this task, participants watch successive short videos in a customized media player. The video content and the order in which videos appear to participants are identical across conditions. I completely disabled the media player controls: participants cannot skip videos (user scrolling is disabled) or use the media player's slider to advance or go back in a particular video.

The conditions differ in their autoplay functionality. In the *Autoplay* condition, videos play automatically when the watching tab is visible and pause with a click anywhere on the video. The media player keeps scrolling to the next videos and continues playing them unless interrupted. In the *Control* condition, participants need to have the watching tab visible and click on each consecutive video to play them. At the end of each video, the scrolling animation brings the next video that only starts playing when the participant clicks on it. Figure 2.2 illustrates the video tab in the *Control* condition.

I manually selected 80 videos to avoid algorithmic biases resulting from the recommender system's personalized suggestions. I curated the videos from YouTube and TikTok, using the tags "funny animal" and "cute animal". I chose animal videos because of their popularity during pretests.[10] I reviewed the selected videos to ensure that they did not contain any harmful or violent actions toward the animals involved. To maximize the frequency of experiencing the autoplay feature during the watching task, I kept the first 60 videos shorter (mean 12.25 seconds, SD 7.04 seconds), with the last 20 videos longer than the rest (mean 25.50 seconds, SD 14.04 seconds).

[10]During the pretests, the experiment was followed by a separate questionnaire asking participants to rate the videos they had seen and suggest what types of videos they would have liked to see more of, with options including sports, news, or trending content.

Figure 2.2: Watching Task tab



Note: This figure describes the "Watch" tab in the *Control* condition. Before the first video is clicked upon to be played, there is a banner that reminds participants that the video won't play without clicking. This banner was not present in the *Autoplay* condition and videos started playing as soon as participants clicked on the "Watch" tab seen on the top right side. Notice that the top part of the following video is already visible in the bottom part of the screen.

2.2.5 Practice Session

The experiment consists of two sessions during which participants interact with the tasks described above. The first session is the practice session on day one. It comes right after the written descriptions of each task and serves to familiarize participants with the tasks and the interface. Participants are assigned to either *Autoplay* or *Control* condition before the practice session, so that they experience the same video interface on both the practice and main sessions.

Practice session was 2 minutes long. Each task was available for only 60 seconds and participants were not allowed to switch between tasks. On the first tab, participants were required to retype CAPTCHAs. The 60-second period was enough to submit one CAPTCHA and see the second CAPTCHA image. After 60 seconds, a pop-up appeared on the screen blocking any other interaction and taking participants to the second page.

Participants were required to close this pop-up to advance to the second task. The rationale behind the pop-up was to ensure participants spent equal time on both tasks. Whether this pop-up was closed was verified and served as an additional attention check.[11] Once closed, the watching task began for 60 seconds, enabling participants to watch 4 to 5 short videos depending on whether they were assigned to the *Autoplay* or the *Control* condition.

2.2.6 Multiple Price List

In the second block of the experiment, I used a Multiple Price List to elicit the Willingness to Pay for commitment against autoplay, commonplace method in the literature (Andersen et al., 2006; Jack et al., 2022). After the practice session on day one, participants made 9 binary decisions, choosing between their preferred media player setting (autoplay off or on) and a bonus payment (see Figure 2.3). I determined bonuses by the per-second earning differences across tasks, multiplied by the expected differences in time spent between the conditions in the first block of the experiment.

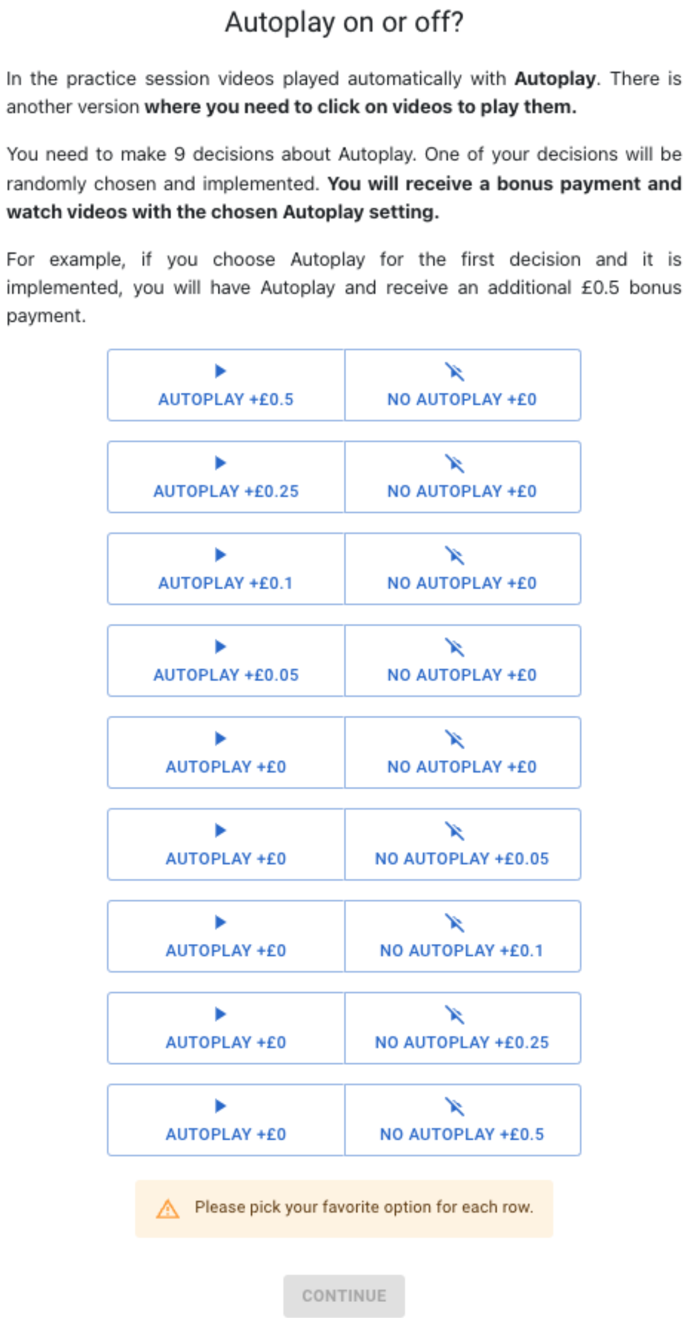
Specifically, I expected the average difference between the *Autoplay* and *Control* conditions, in terms of time spent across tasks, to be in the magnitude of 60 seconds (it was 79.5 seconds, from Table 2.3). This resulted in an average earning difference of $60 * (0.15 - 0.1) = 3$ pence. To capture this in the MPL choices offered, I set the symmetric bonuses accordingly: £[0.05, 0.1, 0.25, 0.5].

I used the Becker-deGroot-Marschak (Becker et al., 1964) random price mechanism for incentivization. Once the binary decisions were completed, I randomly implemented one of their 9 choices and informed participants whether autoplay would be on or off, and

[11] All participants closed this pop-up and switched to the watching task.

of their associated bonus earnings. With this knowledge, participants then made their planned time allocation for day 2.

Figure 2.3: Multiple Price List page



Note: This figure describes the MPL page where participants made 9 binary choices between their preferred autoplay setting and bonus payments. Each row presents a choice between "Autoplay" (left) and "No Autoplay" (right) with associated bonus amounts ranging from £0.05 to £0.5. Note that participants in this block only saw autoplay videos in the practice session, and stated their preferences accordingly. The interface required participants to select one option in each row before proceeding. One of these 9 decisions was randomly selected for implementation in the main session, determining both the autoplay setting and any bonus payment for day two.

2.2.7 Time Choice

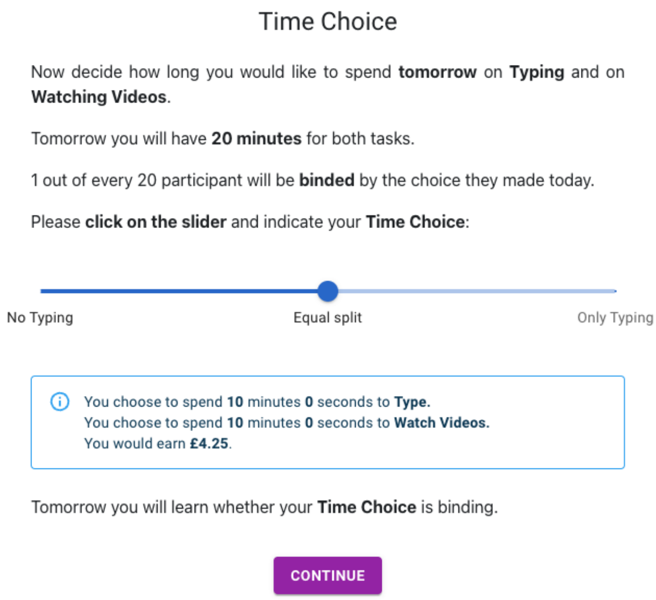
Following the practice session, participants proceeded with the time choice section. This section required them to specify their preferred time allocation between the two tasks for the day 2. Participants were informed the main session would be 20 minutes long, and they were free to choose corner solutions, even if that meant focusing solely on one task.

I designed a slider tool with 20-second steps for participants to indicate their time allocation preferences. Initially, the slider appeared grayed out and required a click to activate. This design aimed to ensure participants considered their choices in an active manner. The click-to-activate design had two benefits: It guaranteed participants interacted with the slider and helped participants understand the task-related payoff structure.

To incentivize this decision, I randomly made 5% of the day 1 choices binding in the main session. Binding choices led to payments based on the initial time choice, but only if participants met the quality standards described under Section 2.2.4 during day 2. I excluded data from these participants from the final dataset accordingly.

The time choice served as a benchmark for participants, and participants accordingly received a reminder of it on day two. The time choice section offers insights into participants' informed time preferences and allows verifying their commitment to these choices after a 24-hour period.

Figure 2.4: Time Choice page



Note: This figure shows the time choice interface with a slider tool with 20-second steps allowing participants to allocate time between the two experimental tasks for the following day's main session. The slider initially appears grayed out and requires a click to activate. Real-time feedback shows the selected allocation (here, 10 minutes each) and expected earnings (£4.25). The interface informs participants that 1 out of 20 (5 percent) will be bound by their choice and that they will learn whether their choice is binding on day two.

2.2.8 Main Session

The main session took place on day two, approximately 24 hours after the time choice. I began by reminding participants of their previous day's time allocation and informing them whether their time choice was binding. 95% of participants were free to allocate their time on day 2 as they preferred.

The main session lasted 20 minutes. Participants could switch between the transcription and watching tasks at any time using the tab interface. Unlike the practice session, there were no time restrictions on individual tasks. Participants had complete autonomy over their time allocation decisions, and could monitor remaining time through the progress bar and the timer. Throughout the main session, I monitored writing and transcribing actions in real-time.

The treatment was implemented through the media player's autoplay functionality as previously described in Section 2.2.4.

2.2.9 Hypotheses

User interface design choices are "not separable from other content choices, since their effect depends on where the platform has positioned its content on the underlying content manifold." (Kleinberg et al., 2022). If content has very low inherent appeal and stickiness, autoplay may have minimal effect on video consumption. An analogy is how introducing breaks would have little impact on engagement for low-stickiness content like documentaries, but dramatically reduce engagement for high-stickiness content like celebrity gossip.[12] Assuming that my selection of videos have sufficient appeal and stickiness, *Autoplay* condition leads to consumption beyond users' stated preferences, creating a measurable gap between planned the time allocation and actual behavior.

Hypothesis 1: *Autoplay* condition decreases the realized time allocated to the transcription task while increasing video consumption. I test this using a linear regression:

$$Y_i = \alpha + \gamma \text{Autoplay}_i + \beta \text{TimeChoice}_i + \delta(\text{Autoplay}_i \times \text{TimeChoice}_i) + \epsilon_i \tag{2.1}$$

where Y_i represents the outcome variables of interest (proportion of time spent on transcription or number of videos watched), Autoplay_i is a binary indicator for the condition, and TimeChoice_i is the standardized proportion of time participants planned to spend on transcription on day 1. I estimate this model with and without the interaction term to examine whether the effect of autoplay varies based on initial time allocation plans. I expect $\gamma < 0$ for transcription and $\gamma > 0$ for videos watched.

Hypothesis 2: *Autoplay* condition has an effect on deviations from stated time allocation preferences (actual minus planned time spent on transcription). I test this using a two-sample t-test comparing the means between the two conditions.

I expect the effect of autoplay on content consumption is known by participants, although they will underestimate how much they would be affected on average (Bongard-Blanchy et

[12]See Bao (2025) for recent work on addictive short drama series.

al., 2021). The demand for a commitment device is driven by "sophisticated" participants, i.e., a subset of participants who are aware of the effect of autoplay, in line with the literature (Ericson & Laibson, 2019).

Hypothesis 3: If participants are aware of their self-control issues, they should exhibit negative Willingness To Pay (WTP) for autoplay, preferring the commitment devices that restrict automatic video transitions. I test this using a logistic regression:

$$Y_{ij} = \beta_0 + \beta_1 \text{Bonus}_{ij} + \epsilon_{ij} \tag{2.2}$$

where Y_{ij} is the binary choice of taking autoplay given the bonus j from the MPL for individual i. I calculate the WTP for autoplay as the sample-averaged effect of bonus $-\beta_0/\beta_1$, and expect a negative sign indicating preference for commitment against autoplay.

In sum, the preregistered hypotheses test whether autoplay distorts optimal time allocation decisions in a setting where participants face monetary incentives favoring the transcription task.

2.3 Sample and Procedures

The experiment was run entirely online, using widely available open-source tools.[13] I sampled from the Prolific participant pool, an online recruitment platform commonly used for academic research, and targeted adult British residents. I further imposed the following criteria: Fluency in English, balanced gender representation, and having a stable internet connection. I did not allow mobile devices and restricted access to the experiment to Chromium-based browsers to enhance the internal validity as previously discussed in Section 2.2.3. Only participants who succeeded on both attention checks, reported not having engaged in other activities, and did not have connection issues during the experiment were included in the final dataset.

Payment was calculated based on actual time spent on each task during the main session on day 2, with compensation rates of 0.15 pence per second for transcription and 0.1 pence per second for watching. Participants needed to meet the established quality

[13] I used the MERN stack to develop and host the website in a free way.

requirements for the transcription task to receive compensation for time spent on that activity. The session concluded automatically after 20 minutes, and participants received their final payment calculation before completing the experiment. Participants received a fixed participation fee of £2.75. The study took less than 30 minutes in total across 2 days, with average earning per participant at £4.08. This amounted to an average payment of £1.33 (median £1.49) for the main session. Participants who completed all required elements received payment within seven days after the second session.[14]

For the first block of the experiment, I preregistered a two-sample parallel design using pretest data and tested for equality of means across conditions. I found that the required sample size to test Hypothesis 1 with the usual parameter values ($\alpha = .05$, $\beta = .2$) would be 114 per condition, with an effect size of 65 seconds calculated by the difference in the average time spent between conditions ($\mu_{Autoplay} - \mu_{Control}$) and population standard deviation of 165 seconds. Since attrition in longitudinal experiments is commonplace, I invited 301 participants on the first day. On day two, I invited the same 301 participants and received 276 complete responses (91% response rate). No participants failed the attention checks more than twice.[15] I dropped 17 participants who self-reported having connection issues and 24 additional participants who stated having engaged in other activities. I also removed 11 participants who had their time choice enforced randomly. Finally, I removed 40 participants that were detected to spend more than 20 consecutive seconds away from our experimental platform who disproportionately come from the *Autoplay* condition. These 40 participants all reported engaging with videos but had their mouse detected outside the experiment window, making it impossible to distinguish whether they genuinely watched videos or browsed other tabs/windows during the study (see Appendix Table A.1). The final dataset on which I base the analysis consisted of 184 individuals. Table 2.2 shows the demographic characteristics of the observed individuals.

[14] Prolific payment policy gives researchers 21 days to transfer funds to participants. The typical delay between study completion and payment is however only 3 to 4 days on average.

[15] Minimum requirement by Prolific attention check policy.

Table 2.2: Balance table for the sample

	Autoplay N = 79	Control N = 105	p [A = C]
Gender			
Female	36 (45.57%)	53 (50.48%)	0.510
Age			
Mean (SD)	42.63 (11.70)	42.03 (11.24)	0.725
Employment			
Employed (full time)	45 (56.96%)	57 (54.29%)	0.718
Employed (part time/self)	23 (29.11%)	25 (23.81%)	0.417
Not employed	11 (13.92%)	23 (21.90%)	0.167
Income			
Low (0-15)	24 (30.38%)	32 (30.48%)	0.989
Middle (15-50)	42 (53.16%)	61 (58.10%)	0.505
High (50+)	13 (16.46%)	12 (11.43%)	0.325
Marital Status			
Married/Partnership	43 (54.43%)	52 (49.52%)	0.510
Single	31 (39.24%)	46 (43.81%)	0.534
Previously married	5 (6.33%)	7 (6.67%)	0.927

Note: This table presents demographic characteristics for the sample of 184 participants across conditions. Income categories are presented in thousands of British pounds. Employment categories: "Not employed" includes unemployed (looking and not looking) and homemaker. "Previously married" includes divorced, widowed, and separated. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests. For employment, part-time and self-employed are together as one category. The sample demonstrates successful randomization with balanced participant characteristics across the conditions.

For the second block of the experiment, I applied the same selection criteria as before and invited 60 participants on day one. On day two, I received 57 complete answers (95% response rate). Once again, no participant failed the attention checks. No participant reported having connection issues. I dropped 5 participants who reported having engaged in other activities. As I was interested in the demand for a commitment device, participants did not have their Time Choice enforced in this treatment. I therefore ended up with 52 individuals in the final MPL dataset (see Appendix Table A.2 for a comparison with the main sample).

2.4 Results

2.4.1 Does Autoplay increase video consumption?

I test my primary hypothesis that autoplay increases video consumption by comparing behavior across conditions. Table 2.3 presents summary statistics for key outcome variables from the 20-minute main session. Two-sample t-tests reveal no significant differences between the *Autoplay* and *Control* conditions.

Table 2.3: Sample statistics for variables of interest

	Autoplay N = 79	Control N = 105	p [A = C]
Time Choice			
Mean (SD)	0.48 (0.31)	0.44 (0.33)	0.410
Median [Min, Max]	0.50 [0.00, 1.00]	0.50 [0.00, 1.00]	
Transcription Proportion			
Mean (SD)	0.65 (0.31)	0.58 (0.33)	0.173
Median [Min, Max]	0.72 [0.00, 1.00]	0.56 [0.00, 1.00]	
Nb. of sessions			
Mean (SD)	5.76 (4.98)	4.89 (5.09)	0.246
Median [Min, Max]	4.00 [1.00, 18.00]	3.00 [1.00, 28.00]	
Session length			
Mean (SD)	456.11 (383.22)	546.35 (411.53)	0.131
Median [Min, Max]	314.75 [68.17, 1212.00]	418.67 [44.93, 1314.00]	
Nb. of CAPTCHA submissions			
Mean (SD)	19.76 (12.11)	17.21 (11.62)	0.150
Median [Min, Max]	19.00 [0.00, 53.00]	15.00 [0.00, 45.00]	
Nb. of videos watched			
Mean (SD)	31.33 (26.34)	32.72 (26.39)	0.723
Median [Min, Max]	25.00 [0.00, 80.00]	34.00 [0.00, 77.00]	
Content rating			
Mean (SD)	6.96 (2.18)	6.62 (2.45)	0.327
Median [Min, Max]	7.00 [2.00, 10.00]	7.00 [0.00, 10.00]	

Note: This table presents key outcome variables from the Main Session. "Time Choice" refers to participants' stated preferences for transcribing on day 1 (proportion), while "Transcription Proportion" represents the actual proportion of time spent on transcription task during the 20-minute Main Session on day 2. Nb. of sessions is the number of transcription or watching sessions participants completed, and Session length represents the mean duration of individual sessions in seconds. Content rating reflects participants' appreciation of video content on a 0-10 scale. p-values are from two-sample t-tests comparing experimental conditions.

To formally test Hypothesis 1, the effects of the *Autoplay* condition on time spent on transcription, I run the regression specified in Equation 4.1. My analysis reveals that the autoplay manipulation had no statistically significant effect on actual transcription behavior. Table 2.4 presents results for transcription proportion as the dependent variable. The coefficient on the autoplay ranges from 0.12 to 0.20 across specifications but remains statistically insignificant in all models (p-values > 0.16). Participants' day 1 time allocations emerge as a powerful predictor of day 2 behavior: a one standard deviation increase in planned transcription time is associated with a 0.67 standard deviation increase in actual transcription time ($p < 0.001$). Including day 1 time allocations increases the variance explained by the model, rising from 1% to 46%.

Table 2.4: *Autoplay* condition and time spent transcribing

	Transcription proportion (z-score)		
	(1)	(2)	(3)
Autoplay	0.203	0.121	0.119
	(0.147)	(0.109)	(0.110)
Time Choice (z-score)		0.673***	0.651***
		(0.057)	(0.081)
Autoplay × Time Choice			0.057
			(0.110)
Constant	−0.087	−0.052	−0.053
	(0.100)	(0.076)	(0.076)
Obs.	184	184	184
R^2	0.010	0.462	0.463
F-statistic	1.91	77.42	61.52

Note: Robust standard errors in parentheses. All variables are standardized. Time choice represents participants' planned transcription proportion from day 1. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Similarly, Table 2.5 examines the effect on video consumption, a direct test of my primary hypothesis that autoplay would increase video consumption. Across all specifications, the autoplay coefficient is statistically insignificant. This suggests that the autoplay feature did not meaningfully alter participants' video consumption behavior. Again, day 1 plans prove to be the dominant predictor: participants who planned more transcribing watched

−0.66 SD fewer videos ($p < 0.001$).

Table 2.5: *Autoplay* condition and the number of videos watched

	Nb. of Videos Watched (z-score)		
	(1)	(2)	(3)
Autoplay	−0.053	0.029	0.031
	(0.149)	(0.112)	(0.112)
Time Choice (z-score)		−0.664***	−0.619***
		(0.056)	(0.079)
Autoplay × Time Choice			−0.114
			(0.105)
Constant	0.023	−0.012	−0.010
	(0.098)	(0.076)	(0.076)
Obs.	184	184	184
R^2	0.001	0.440	0.443
F-statistic	0.13	73.06	61.53

Note: Robust standard errors in parentheses. All variables are standardized. Time choice represents participants' planned transcription proportion from day 1. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

For completeness, I provide the regression outputs from the unrestricted dataset in Appendix Table A.3 and A.4. These include 40 participants who spent at least 20 consecutive seconds outside either task. Comparing the interaction models (column 3), I observe that the effect of the *Autoplay* condition on transcribing behavior decreases from 0.119 SD in the restricted sample to 0.055 SD in the unrestricted sample. The effect of the *Autoplay* condition on the number of videos watched increases from 0.031 SD in the restricted sample to 0.132 SD in the unrestricted sample. However, the null results remain unchanged across both specifications. Because I cannot ascertain the data quality for the unrestricted sample, I base my conclusions on the restricted sample of 184.

Two key patterns emerge that help explain the null findings. First, the contradicting results: the *Autoplay* condition increases time spent transcribing by 0.12 standard deviations (though statistically insignificant) while simultaneously increasing the number of videos watched by 0.03 standard deviations (also insignificant). If people type more, shouldn't they watch fewer videos? I identify a mechanical confound in my experimental design that explains this result: the time lost during video transitions. Participants in the *Control* condition, who had to manually click to start each new video, lost an average of 46.52 seconds due to these transition delays whereas those in the *Autoplay* condition lost only 1.08 seconds. This large difference of 45.45 seconds per session is statistically significant ($t(105) = 4.89$, $p < 0.001$). This difference in time lost not watching videos explains the contradicting results and underscores the importance of looking at the unbiased outcome variable (i.e., number of videos watched) when evaluating the effects of the autoplay feature.

Second, assuming the actual effect size for the number of videos watched lies somewhere between the restricted and unrestricted samples, the power calculations based on pretests were insufficient: The required sample size to provide a two-sided test for Hypothesis 1b with the usual parameter values ($\alpha = .05$, $\beta = .2$) would be between 300 and 5000 per treatment, with common standard deviation as the root MSE from the standardized number of videos watched variable in the third regression. Other solutions to this power problem are design-related and involve extending the main-session duration to above 20 minutes, or nudging participants to mix more between tasks to avoid corner solutions (e.g., higher earnings for time allocations between .25 and .75 of either task).

Table 2.6: Seconds lost by condition

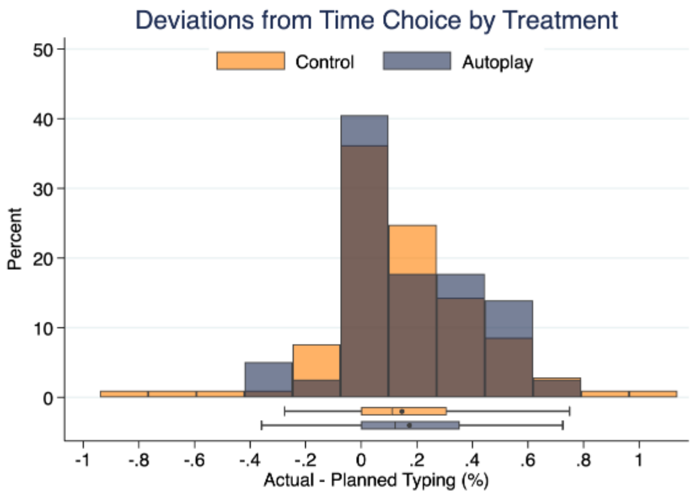
	Control	Autoplay	Difference
Total seconds lost	46.52	1.08	45.45***
	(9.27)	(0.65)	(9.29)
Observations	105	79	184
t-statistic			4.89

Note: Standard errors are in parentheses. Time is measured as the average of seconds lost to video transitions per session. The difference was tested using a two-sample t-test with unequal variances. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

2.4.2 Autoplay and deviations from day 1 time allocation

Does the *Autoplay* condition have an effect on deviations from day 1 time choice? To assess Hypothesis 2, I calculated the deviation by subtracting participants' planned transcribing proportion from their actual transcribing proportion (actual minus planned). Figure 2.5 shows the distribution of deviations from time choice across conditions, which appears right-skewed with positive values indicating that participants transcribed more than their day 1 allocation. Participants in the *Control* condition deviated by an average of 14.6 percentage points, while those in the *Autoplay* condition deviated by 17.3 percentage points. A two-sample t-test reveals no statistically significant difference between means ($t = -0.694$, $p = 0.488$). The Kolmogorov-Smirnov test similarly finds no significant difference in distributions across conditions ($D = 0.085$, $p = 0.903$). These results indicate that participants in both conditions spent more time on the transcription task than they initially planned. However, the *Autoplay* condition did not significantly alter the magnitude or pattern of these deviations. Contrary to my hypothesis, I find no evidence that *Autoplay* condition led participants to deviate toward consuming more videos.

Figure 2.5: Distribution of deviations from day 1 time allocation by condition



Note: This figure shows the distribution of deviations from day 1 time allocation plans, calculated as actual transcribing proportion minus planned transcribing proportion. Positive values indicate participants transcribed more than they planned. The distributions are similar across conditions, and show over-transcribing relative to initial plans.

2.4.3 Willingness To Pay for autoplay

Contrary to Hypothesis 3, participants exhibited positive WTP for autoplay rather than demand for commitment devices against it. I assessed the demand for commitment by comparing the share of participants choosing to turn autoplay on or off along with the associated bonus payments for each of the nine decisions in the price list. To quantify participants' WTP for autoplay, I estimate the regression specification in Equation 2.2, where the dependent variable is the binary choice of selecting autoplay (1) or turning it off (0), and the independent variable is the bonus amount in pounds from the MPL.

Results are presented in Table 2.7. The coefficient on the bonus amount is 27.97 ($p < 0.001$), indicating that participants are sensitive to the financial cost of keeping autoplay on. The WTP for autoplay is the negative ratio of the constant to the bonus coefficient: $-0.6263/27.9729 = -0.0224$ pounds, or 2.24 pence for the 20-minute session. On average, participants value autoplay at 6.72 pence/hour and would be willing to forgo this amount to maintain access to the feature rather than have it turned off. Participants in this study did not appear to perceive a tradeoff from autoplay in increasing video consumption.

Table 2.7: Willingness to Pay for autoplay

	(1)
Bonus (£)	27.973***
	(6.277)
Constant	0.626**
	(0.199)
Obs.	468
Clusters	52
Pseudo R^2	0.658
Wald $\chi^2(1)$	19.86

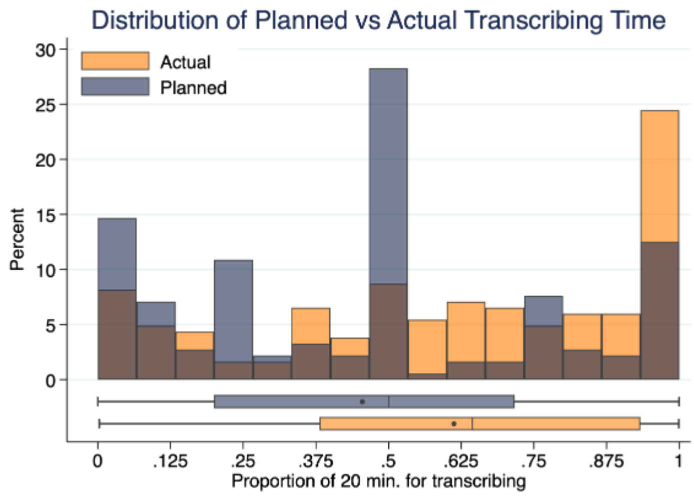
Note: Logistic regression with standard errors clustered at the participant level in parentheses. Dependent variable is binary choice of selecting autoplay (1) or turning it off (0). Bonus amounts range from £-0.5 to £0.5. The pseudo R^2 indicates good fit. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

2.4.4 Understanding the Null Results

Planned versus Actual Time Allocation

Participants allocated more time to transcription than initially planned. On average, participants planned to spend 45.5% of their time transcribing but actually spent 61.2%, representing an increase of 15.7 percentage points ($t = 8.24$, $p < 0.001$). The entire distribution shifted rightward (see Figure 2.6), with median actual transcription time (64.3%) substantially exceeding median planned time (50.0%).

Figure 2.6: Distribution of planned vs actual time allocation



Note: This figure compares the distribution of participants' planned time allocation (Day 1) with their actual time spent transcribing (Day 2). The rightward shift demonstrates participants' systematic tendency to transcribe more than originally planned.

The pattern is more pronounced when examining individual transitions. Among 151 participants who planned to spend less than three-quarters of their time transcribing, 106 (70%) actually transcribed for three-quarters or more of the session. In contrast, among 33 participants who planned to spend three-quarters or more of their time transcribing, only 4 (12%) reduced below 75%. This asymmetric upward shift contradicts traditional models of time inconsistency, which predict succumbing to immediate temptation rather than increasing effortful work.

Table 2.8: Transition matrix: Planned vs actual time allocation

	Planned (Day 1)	
Actual (Day 2)	<75% transcription	≥75% transcription
<75% transcription	45 (30%)	29 (88%)
≥75% transcription	106 (70%)	4 (12%)
Total	151	33

Note: This table shows transitions between planned and actual time allocation patterns. The asymmetric pattern demonstrates systematic shift toward transcription, with 106 participants increasing transcription time versus only 4 decreasing it.

The deviation toward the more effortful transcription task is rather unexpected, and participant perceptions of the decision-making environment may have influenced their behavior. This result goes against my hypothesis that participants would deviate toward the tempting video content, and points to the need to examine it more closely.

Experimenter demand effects

Why did participants end up transcribing more than their day 1 plans? To understand how participants perceived the decision-making environment, I asked how they decided to allocate their time during the main session. I analyzed participants' qualitative responses to this open-ended question using dictionary-based concept detection methods (Ash & Hansen, 2023).

I identified words related to negative perceptions of the transcription task (e.g., "tiresome", "frustrating", "difficult") and break-taking behavior when switching to watching (e.g., "break", "rest", "relax"). Figures 2.7a and 2.7b show the frequency of these concepts in participants' responses.

Figure 2.7: Word clouds for participant perceptions

(a) Typing task negative descriptions

(b) Changing to watching task

Note: Panel (a) shows words describing negative perceptions of the transcription task. Panel (b) shows words describing switching to the watching task as break-taking behavior. Word size corresponds to frequency of mention across responses.

When asked to describe their allocation decisions, 29% of participants described the transcription task negatively, while 23% specifically described watching videos as taking breaks from transcription. These findings suggest that participants perceived the experiment as a work environment where they should transcribe as much as possible while taking breaks to watch the videos. Combined with the systematic increases in transcription time, experimenter demand effects have influenced participant behavior as suspected.

Effect of achieving day 1 time allocation

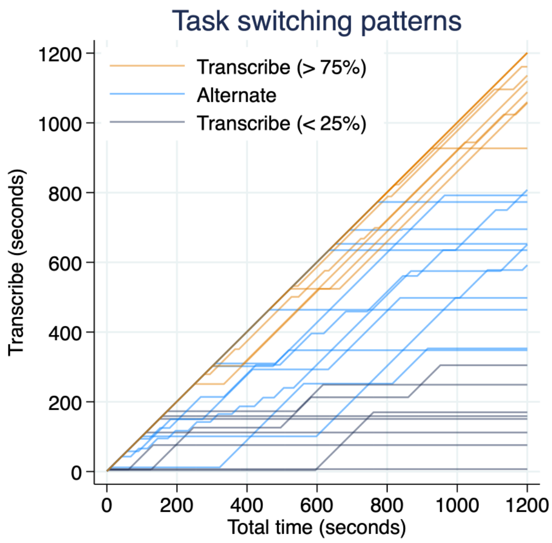
If experimenter demand effects are driving participants to transcribe more than planned, participants may ignore their day 1 goals entirely and continue transcribing regardless of achieving their stated preferences. Alternatively, if participants genuinely care about their day 1 plans, we should observe behavioral changes when they reach their planned transcribing duration—specifically, a reduction in transcribing effort after achieving their goal.

I examine this by analyzing task switching patterns throughout the 20-minute session.

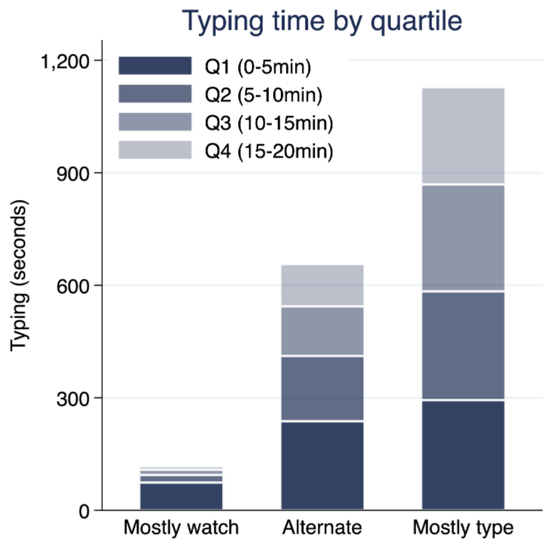
Figure 2.8a shows substantial heterogeneity in individual transcribing patterns, with transcribing periods interrupted by watching periods for participants who alternate between tasks. Dividing the session into 5-minute quarters reveals that transcribing periods become progressively shorter over time (Figure 2.8b), suggesting that task switching may be strategic as participants either achieve their day 1 goals or experience fatigue.

Figure 2.8: Time spent transcribing

(a) Individual transcribing patterns



(b) Average transcribing by quartile



Note: Panel (a) shows transcribing time for 30 random participants across the 20-minute session. Panel (b) shows average transcribing time by 5-minute periods, revealing declining transcribing duration over time.

To test whether participants use their day 1 plans as behavioral reference points, I implement a sharp regression discontinuity design using the moment participants reach their planned transcribing duration as the cutoff. I estimate the discontinuous change in the share of participants choosing to transcribe task at this threshold.

Results in Table 2.9 show a significant effect: participants are 27.4 percentage points less likely to continue transcribing in the 30-second window after achieving their plan ($p = 0.007$) and 25.5 percentage points less likely in the 60-second window ($p = 0.003$). This effect dissipates by the 90-second window, becoming statistically insignificant. These findings indicate that participants do treat plan completion as a behavioral milestone, suggesting that experimenter demand effects were not so strong as to completely override their day 1 preferences.

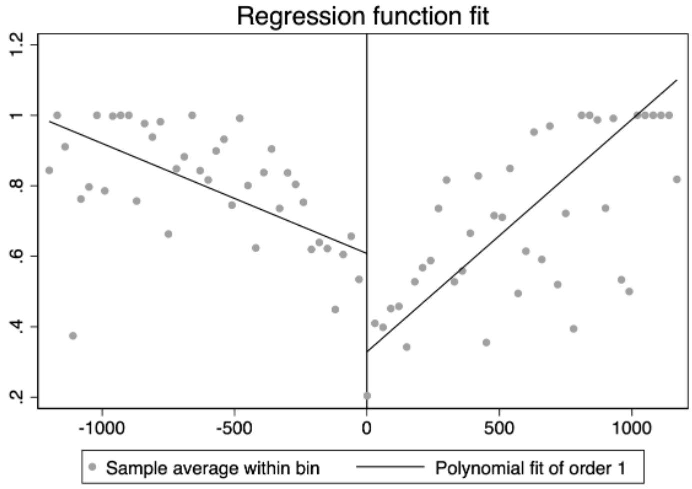
Table 2.9: Regression Discontinuity estimates on transcribing after achieving day 1 plans

	Time Window		
	30-seconds	60-seconds	90-seconds
Proportion transcribing	-0.274***	-0.255***	-0.075
	(0.101)	(0.087)	(0.108)
Obs. (left)	66,786	67,985	67,119
Obs. (right)	80,270	89,006	92,999
Clusters (left)	157	154	155
Clusters (right)	154	157	161
Bandwidth	339.9	368.3	422.3

Note: This table presents robust regression discontinuity estimates of the treatment effect on task choice at the moment participants achieve their day 1 planned transcribing duration. Standard errors clustered by participant in parentheses. The outcome variable is a binary indicator for choosing the transcribing task. Running variables are binned at 30-, 60-, and 90-second intervals around the achievement threshold. All estimates use triangular kernel with MSE-optimal bandwidth selection. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

However, the discontinuity analysis reveals additional complexity. Figure 2.9 shows that while participants initially reduce transcribing after achieving their day 1 plans, they subsequently increase transcribing effort again. This pattern suggests that participants have indeed found the watching task less engaging than anticipated, leading them to return to transcribing despite having already met their stated preferences.

Figure 2.9: Linear fit for transcribing rate before and after achieving day 1 plan



Note: This figure plots the average transcribing rate across 30-second bins for participants as they approach their day 1 goal. The negative slope on the left-hand side indicates participants working less and less as they approach the cutoff. The drop showcases the discontinuity where achieving the day 1 goal causes in transcribing behavior. The positive slope on the right-hand side reveals another trend: As time goes on, participants tend to go back to the transcribing task and decrease their video consumption. This demonstrates participants' tendency to spend more time on the effortful transcribing task than originally planned.

2.5 Conclusion

This study examined whether autoplay features in isolation can override users' stated preferences for media consumption. In a two-day online experiment with 236 participants, we find no evidence that autoplay increases video consumption when content is held constant. Participants viewed autoplay favorably, exhibiting positive willingness-to-pay for the convenience it provides rather than perceiving it as a self-control problem. However, experimenter demand effects resulted in participants allocating more time to transcription than planned, confounding the intended autoplay treatment.

These findings highlight a fundamental challenge of studying potentially addictive design features in controlled laboratory settings. Design features like autoplay appear to derive their persuasive power from personalized content streams and algorithmic curation. In my controlled setting with identical, manually curated videos, I likely stripped away the very elements that make autoplay compelling in practice.

These results have implications for both research methodology and policy. Laboratory studies that isolate interface features from their natural digital ecosystem may fail to capture their real-world behavioral consequences. The controlled conditions necessary for internal validity inadvertently eliminate the personalized content that drives engagement on actual platforms. While autoplay shows no effect in my experimental setting, field studies suggest stronger impacts in natural usage contexts (Hiniker et al., 2018; Schaffner et al., 2025). Future research should prioritize better approximating genuine media consumption environments in the field (see e.g., Beknazar-Yuzbashev et al. (2025); Purohit et al. (2023)) while maintaining the control necessary for causal inference.

References

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–76.

Allcott, H., Gentzkow, M., & Song, L. (2022). Digital addiction. *American Economic Review*, *112*(7), 2424–2463.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006, December). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383–405. doi: 10.1007/s10683-006-7055-6

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, *21*(1), 99–131. doi: 10.1007/s10683-017-9527-2

Ash, E., & Hansen, S. (2023). Text Algorithms in Economics. *Annual Review of Economics*, *15*(Volume 15, 2023), 659–688. doi: 10.1146/annurev-economics-082222-074352

Bao, R. (2025). *"just one more clip": Short videos, big self-control problems* [Working Paper].

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. doi: 10.1002/bs.3830090304

Beknazar-Yuzbashev, G., Jiménez-Durán, R., McCrosky, J., & Stalinski, M. (2025). *Toxic content and user engagement on social media: Evidence from a field experiment* [CESifo Working Paper].

Bhatia, S., Crawford, M. M., McDonald, R. L., Moreno, M. A., & Read, D. (2021). *Inconsistent Planning and the Allocation of Tasks Over Time* (OSF Preprints No. b4mg7). Center for Open Science.

Bonein, A., & Denant-Boèmont, L. (2015). Self-control, commitment and peer pressure: A laboratory experiment. *Experimental Economics*, *18*(4), 543–568. doi: 10.1007/s10683-014-9419-7

Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzini, G. (2021). " i am definitely manipulated, even when i am aware of it. it's ridiculous!"-dark patterns from the end-user perspective. In *Designing interactive systems conference 2021* (pp. 763–776).

Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American*

Economic Review, *112*(11), 3660–3693.

Bursztyn, L., Handel, B. R., Jimenez, R., & Roth, C. (2023). *When Product Markets Become Collective Traps: The Case of Social Media* (Working Paper No. 31771). National Bureau of Economic Research. doi: 10.3386/w31771

Chen, Y., Fu, Y., Chen, Z., Radesky, J., & Hiniker, A. (2025). *The Engagement-Prolonging Designs Teens Encounter on Very Large Online Platforms* (No. arXiv:2411.12083). arXiv. doi: 10.48550/arXiv.2411.12083

Collis, A., & Eggers, F. (2022). Effects of restricting social media usage on wellbeing and performance: A randomized control trial among students. *PloS one*, *17*(8), e0272416.

Ek, C., & Samahita, M. (2023). Too much commitment? an online experiment with tempting youtube content. *Journal of Economic Behavior & Organization*, *208*, 21–38.

Ericson, K. M., & Laibson, D. (2019). Intertemporal choice. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 1–67). Elsevier. doi: 10.1016/bs.hesbe.2018.12.001

Groshek, J., Krongard, S., & Zhang, Y. (2018). Netflix and Ill? Emotional and Health Implications of Binge Watching Streaming TV. In *Proceedings of the 9th International Conference on Social Media and Society* (pp. 296–300). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3217804.3217932

Hiniker, A., Heung, S. S., Hong, S., & Kientz, J. A. (2018). Coco's videos: an empirical investigation of video-player design features and children's media use. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–13).

Hoong, R. (2021). Self control and smartphone use: An experimental study of soft commitment devices. *European Economic Review*, *140*, 103924. doi: 10.1016/j.euroecorev.2021.103924

Houser, D., Schunk, D., Winter, J., & Xiao, E. (2018). Temptation and commitment in the laboratory. *Games and Economic Behavior*, *107*, 329–344. doi: 10.1016/j.geb.2017.10.025

Jack, B. K., McDermott, K., & Sautmann, A. (2022, November). Multiple price lists for willingness to pay elicitation. *Journal of Development Economics*, *159*, 102977. doi: 10.1016/j.jdeveco.2022.102977

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2022). The challenge of understanding what users want: Inconsistent preferences and engagement optimization. *arXiv preprint arXiv:2202.11776*.

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of experimental psychology. General*, *143*(1), 131.

Lukoff, K., Lyngs, U., Zade, H., Liao, J. V., Choi, J., Fan, K., ... Hiniker, A. (2021). How the design of youtube influences user sense of agency. In *Proceedings of the 2021 chi conference on human factors in computing systems* (pp. 1–17).

Lupiáñez-Villanueva, W. F., Boluda, A., Bogliacino, F., & Liva, G. (2022). Behavioural study on unfair commercial practices in the digital environment: Dark patterns and manipulative personalisation. , 303.

Lyngs, U., Lukoff, K., Slovak, P., Binns, R., Slack, A., Inzlicht, M., ... Shadbolt, N. (2019). Self-Control in Cyberspace: Applying Dual Systems Theory to a Review of Digital Self-Control Tools. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3290605.3300361

Lyngs, U., Lukoff, K., Slovak, P., Seymour, W., Webb, H., Jirotka, M., ... Shadbolt, N. (2020). 'I Just Want to Hack Myself to Not Get Distracted': Evaluating Design Interventions for Self-Control on Facebook. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3313831.3376672

Mathur, A., Kshirsagar, M., & Mayer, J. (2021). What Makes a Dark Pattern... Dark?: Design Attributes, Normative Considerations, and Measurement Methods. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). Yokohama Japan: ACM. doi: 10.1145/3411764.3445610

Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., ... others (2023). Like-minded sources on facebook are prevalent but not polarizing. *Nature*, *620*(7972), 137–144.

Purohit, A. K., Bergram, K., Barclay, L., Bezençon, V., & Holzer, A. (2023). Starving the newsfeed for social media detox: Effects of strict and self-regulated facebook newsfeed diets. In *Proceedings of the 2023 chi conference on human factors in computing systems* (pp. 1–16).

Schaffner, B., Stefanescu, A., Campili, O., & Chetty, M. (2023). Don't Let Netflix Drive the Bus: User's Sense of Agency Over Time and Content Choice on Netflix. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), 1–32. doi: 10.1145/3579604

Schaffner, B., Ulloa, Y., Sahni, R., Li, J., Cohen, A. K., Messier, N., ... Chetty, M. (2025). An Experimental Study Of Netflix Use and the Effects of Autoplay on Watching Behaviors. *Proc. ACM Hum.-Comput. Interact.*, *9*(2), CSCW030:1–CSCW030:22. doi: 10.1145/3710928

Silverman, J., Srna, S., & Etkin, J. (2024). *Can Time Limits Increase Time Spent?* (SSRN Scholarly Paper No. 4381779). Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.4381779

Walton-Pattison, E., Dombrowski, S. U., & Presseau, J. (2018). 'Just one more episode': Frequency and theoretical correlates of television binge watching. *Journal of Health Psychology*, *23*(1), 17–24. doi: 10.1177/1359105316643379

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98. doi: 10.1007/s10683-009-9230-z

A.1 Additional Figures and Tables

Additional Tables

Table A.1: Comparison of main sample vs flagged participants

	Main Sample N = 184	Flagged N = 40	p [M = F]
Gender			
Female	89 (48.37%)	13 (32.50%)	0.068*
Age			
Mean (SD)	42.29 (11.41)	42.85 (10.75)	0.768
Employment			
Employed (full time)	102 (55.43%)	27 (67.50%)	0.162
Employed (part time/self)	48 (26.09%)	8 (20.00%)	0.420
Not employed	34 (18.48%)	5 (12.50%)	0.366
Income			
Low (0-15)	56 (30.43%)	6 (15.00%)	0.048**
Middle (15-50)	103 (55.98%)	31 (77.50%)	0.012**
High (50+)	25 (13.59%)	3 (7.50%)	0.291
Marital Status			
Married/Partnership	95 (51.63%)	21 (52.50%)	0.921
Single	77 (41.85%)	17 (42.50%)	0.940
Previously married	12 (6.52%)	2 (5.00%)	0.719
Treatment Assignment			
Autoplay condition	79 (42.93%)	25 (62.50%)	0.025**

Note: This table compares demographic characteristics between the main sample (N=184) and participants flagged for potential inattention. Flagged participants reported engaging with videos but had mouse activity detected outside the experiment window, making it impossible to verify genuine engagement. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Comparison of MPL sample vs main sample

	Main Sample N = 184	MPL N = 52	p [M = MPL]
Gender			
Female	89 (48.37%)	26 (50.00%)	0.835
Age			
Mean (SD)	42.29 (11.41)	41.90 (14.14)	0.858
Employment			
Employed (full time)	102 (55.43%)	32 (61.54%)	0.433
Employed (part time/self)	48 (26.09%)	10 (19.23%)	0.311
Not employed	34 (18.48%)	10 (19.23%)	0.902
Income			
Low (0-15)	56 (30.43%)	15 (28.85%)	0.825
Middle (15-50)	103 (55.98%)	30 (57.69%)	0.826
High (50+)	25 (13.59%)	7 (13.46%)	0.981
Marital Status			
Married/Partnership	95 (51.63%)	28 (53.85%)	0.778
Single	77 (41.85%)	19 (36.54%)	0.491
Previously married	12 (6.52%)	5 (9.62%)	0.446

Note: This table compares the demographic characteristics for the main sample (N=184) and the final MPL sample in block 2 (N=52). We observe no significant differences across the two samples.

Table A.3: *Autoplay* condition and time spent transcribing

	Transcription proportion (z-score)		
	(1)	(2)	(3)
Autoplay	0.136	0.055	0.055
	(0.133)	(0.108)	(0.108)
Time Choice (z-score)		0.595***	0.580***
		(0.061)	(0.089)
Autoplay $\times$ Time Choice			0.034
			(0.119)
Constant	-0.063	-0.026	-0.027
	(0.094)	(0.078)	(0.079)
Obs.	224	224	224
R^2	0.005	0.357	0.357
F-statistic	1.04	50.28	35.00

Note: Robust standard errors in parentheses. All variables are standardized. Time choice represents participants' planned transcription proportion from day 1. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table A.4: *Autoplay* condition and number of videos watched

	Nb. of Videos Watched (z-score)		
	(1)	(2)	(3)
Autoplay	0.051	0.131	0.132
	(0.135)	(0.109)	(0.109)
Time Choice (z-score)		-0.590^{***}	-0.555^{***}
		(0.060)	(0.084)
Autoplay × Time Choice			-0.077
			(0.118)
Constant	-0.023	-0.061	-0.059
	(0.090)	(0.075)	(0.075)
Obs.	224	224	224
R^2	0.001	0.347	0.348
F-statistic	0.14	48.91	34.21

Note: Robust standard errors in parentheses. All variables are standardized. Time Choice represents participants' planned typing proportion from day 1. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Chapter 3

Peer skill identification and social class: Evidence from a referral experiment

Status: Not sent out for publication

Contribution Statement: Jhon Díaz: Resources, Investigation. Manu Munoz: Software, Conceptualization, Methodology, Supervision, Funding acquisition, Writing - Review & Editing, Project administration. Ernesto Reuben: Conceptualization, Methodology, Supervision, Funding acquisition, Project administration. Reha Tuncer: Conceptualization, Methodology, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization.

Abstract

Cognitive and social skills are both increasingly valued in the labor market, but social skills are difficult to observe. In the absence of observable signals, peer assessments can be valuable screening tools. We study how well individuals identify productive peers across cognitive and social skills in a lab-in-the-field experiment with 849 university students. After students interact for an entire term, we collect incentivized skill measures from all classmates. We then ask for referrals of the highest scoring peers in each skill, incentivizing referrals based on the nominee's score. To examine potential social class barriers in referrals, we randomly assign half of the participants to receive additional incentives for identifying high-skilled peers from low-socioeconomic status. We find that peers can successfully identify cognitive skills but not social skills of their classmates. There is only evidence of a bias against low-SES peers in unique cognitive skill referrals, and the treatment incentives helps mitigate it. Our findings suggest that the accuracy of peer assessments varies substantially across skill dimensions and appropriate changes in the incentivization structure can make peer assessments robust to existing biases.

JEL Classification: C93, D03, D83, J24

Keywords: network homophily, labor market, performance evaluation, hiring screening, human capital, incentive mechanisms, workplace diversity, academic performance, socioeconomic barriers, information asymmetry

3.1 Introduction

Evaluating the productivity of others is a standard feature of the labor market. Employers assess job candidates, managers evaluate workers for promotion, and team leaders select collaborators based on beliefs about others' capacity to perform well in different tasks. Whenever observable productivity signals such as test scores or past experience are available, decision-makers rely on those to make accurate evaluations. But such signals are scarce for tasks that are interpersonal in their nature and difficult to quantify. In these settings, peer assessments akin to referrals can be a particularly strong screening tool which combines cost-efficiency and accuracy, as sustained interactions among people who work together provide opportunities to directly observe each other's productive qualities in various domains.

However, identifying productive peers across a multitude of productivity dimensions is not straightforward. First, peers could accurately assess productivity in one dimension, but they may struggle to evaluate it in another because of its harder to observe nature. Cognitive and social (interpersonal) skills are two such dimensions of human capital that are increasingly rewarded in the labor market (Deming, 2017, 2023). Second, biases in productivity beliefs can lead to systematic deviations in assessment accuracy. The case for low-socioeconomic status (low-SES) individuals is particularly concerning, as peers may systematically underestimate their abilities due to stereotypes or lack of information. Such biased assessments could contribute to their worse labor market outcomes despite having the necessary skills (Stansbury & Rodriguez, 2024).

The overall purpose of this paper is twofold: To evaluate how accurately peers identify productive others in cognitive and social skills, and whether disadvantaged low-SES individuals face barriers in selection when peers assess productivity across these skills.

We conducted a lab-in-the-field experiment in a Colombian university to answer these questions. After interacting for an entire term (about 4 months) in small classrooms (average 26 students per class), we collected incentivized cognitive and social skill measures from all participants to obtain objective productivity distributions. Participants then assessed classmates' productivity across these dimensions by making referrals, allowing us to compare referred peers to those who were not. We incentivized referrals by bonuses contingent on the nominee's score in the skill measures. Nominees did not receive any benefit from being referred. Both features allowed us to rule out concerns of potential social transfers (i.e., nepotism or favoritism) and reputational costs typical in the referral literature (see for example Bandiera et al. (2009); Witte (2021)). Once we abstracted away from these elements, the referral decision became one of measuring productivity beliefs through nominated candidates.

Even in an incentivized setting like ours, biases about low-SES individuals could be at play because of the underlying beliefs classmates hold about their productivity. To address this we designed two treatments. In the **Baseline** treatment, we gave pure performance incentives to referrals regardless of social class. Participants in the **Quota** treatment received additional incentives to identify high-skilled low-SES peers. To be able to make comparisons within the same referral choice sets, we assigned half of the participants within each classroom to either treatment. This setup allows us to assess how well incentives mitigate the said biases in peer productivity beliefs across the different referral behaviors that we observe.

Our first goal is understanding how well peers identify cognitive and social skills of their classmates under pure performance incentives at **Baseline**. We find that peers have distinct screening abilities for skills, and use different types of referral strategies because of it. Specifically, peers successfully identify cognitive skill but not social skill of their classmates. They also frequently refer the same peers for both skills, at rates much higher than the actual overlap between those who are productive at both cognitive and social skill. For this reason we separately analyzed the three referral types: Those made in common for both skills, and those made uniquely for cognitive or social skill. Common referrals for both skills identified classmates with higher grades but not higher skills. This suggests an observable proxy such as academic performance influences peer productivity assessments in the absence of credible skill information. For unique cognitive skill referrals, both grades and measured cognitive skill are equally good predictors. Unique social skill referrals are not predicted by either academic performance or social skill, suggesting that social skills might be less observable in classroom settings or require different measures to evaluate accurately. These findings reveal a nuanced picture of how peer assessments of productivity may depend on how discernible the skill in question is, and how they can be influenced by the availability of other observable proxies for productivity.

We find limited support for a bias affecting low-SES individuals. Of the three referral types, we find bias only in unique cognitive skill referrals when accounting for peer skills. This characterizes the decisions of about 75% of participants who made at least one unique cognitive skill referral, and about half of all cognitive skill referrals overall. The **Quota** treatment mitigates the bias for this subset of referrals, while not changing the referral rates of low-SES individuals for the rest of the referral strategies that were not biased in the first place. There is also no meaningful efficiency-equity tradeoff affecting productivity of peers referred in the **Quota** treatment. Our findings show peer productivity assessments are robust to salient differences between social classes, and provide evidence that existing biases can be remedied with changes in the incentivization structure without compirimising productivity.

Our paper contributes to various strands of the literature. First, we contribute the literature on referral experiments that strives to understand how referrals help screeening for productive workers. Past work provides causal evidence that peer productivity assessments using referrals bring in productive workers (Pallais & Sands, 2016), and that performance-contingent incentives lead to improvements in the productivity of referred candidates (Beaman et al., 2018; Beaman & Magruder, 2012). These studies allow referrals to be made from different candidate pools where referrers are free to nominate any candidate, and as a result confound screening ability with advantages arising from access to different candidate pools (Montgomery, 1991). We implement common choice sets for referrals which allow us to isolate peers' true screening ability and enable straightforward comparison between experimental treatments in terms of referral choice sets. Our paper complements the literature on referral experiments by providing causal evidence that peers have skill-dependent screening abilities that go beyond the differences in candidate pools under performance-contingent incentives.

Second, we contribute to the growing body of work on the relevance of noncognitive skills in the labor market. This literature examines dimensions of human capital such as patience, self-control, conscientiousness, teamwork, and critical thinking that contribute positively to labor market returns (Heckman & Kautz, 2012; Heckman et al., 2006; Lindqvist & Vestman, 2011; Weinberger, 2014). Among these, interpersonal skills are exceptionally relevant for labor market gains in the last two decades as a complement to cognitive skill (Deming, 2017, 2023). Yet, hiring firms report difficulties in assessing social skills in candidates, and applicants are willing to pay substantial sums to convey social skill feedback to employers (Bassi & Nansamba, 2022). We contribute to this literature with our peer productivity assessments across two dimensions of skills, and show that peers can identify cognitive skill but struggle to assess social skills. Our results suggest that referrals may be ineffective for screening attributes that are less visible or harder to proxy through standard productivity measures in the assessment environment.

Finally, we contribute to the literature on diversity considerations in referrals. Homophily[1] in referrals drives correlations among social groups' employment and wages (Calvo-Armengol & Jackson, 2004; Calvó-Armengol & Jackson, 2007), as individuals are more often tied to others with comparable socioeconomic status (Chetty et al., 2022b). Limited interaction across social classes due to spatial segregation is shown to drive at least some of the differences (Chetty et al., 2022a). In this context, efficiency of diversity treatments in endogenous networks may be constrained by availability. To counter this, we consider a socially diverse university setting where we use exogenously imposed networks, and required participants to refer among classmates. Anticipating differences in referral outcomes for low-SES individuals even when networks across social classes overlap by design, we introduced quota-like incentives as a treatment arm to increase referrals to low-SES peers.[2] Our findings complement the literature on biases in referrals (Beugnot & Peterlé, 2020; Hederos et al., 2024) by first showing the existence of a social class bias and then providing the causal evidence for targeted incentives that effectively reduce the bias in our setting without compirimising productivity.

The remainder of the paper is organized as follows. Section 3.2 begins with the background and setting in Colombia. In Section 3.3 we present the design of the experiment, including the skill assessment, referral and guessing tasks. In Section 3.4 we describe the data and procedures. Section 3.5 discusses the results of the experiment. Section 3.6

[1] A well-documented empirical consistency in sociology where individuals form ties more often with others who are similar to themselves across observable characteristics (McPherson et al., 2006, 2001).

[2] We design the treatment incentives in inspiration from the success of gender quotas in the affirmative action literature (e.g., Balafoutas and Sutter (2012); Bertrand et al. (2019); Niederle et al. (2013)).

concludes. The Appendix presents additional tables and figures as well as the experiment instructions.

3.2 Background and Setting

Our study takes place at a medium-sized private university in Colombia, with approximately 6,000 enrolled students. The university's student body is remarkably diverse with slightly more than half of the students classified as low-SES. This diversity provides a unique research setting, as Colombian society is highly unequal and generally characterized by limited interaction between social classes, with different socioeconomic groups separated by education and geographic residence.[3] Despite significant financial barriers, many lower middle-class families prioritize university education for their children (Hudson & Library of Congress, 2010, p. 103), with our partner institution representing one of the few environments where sustained inter-class contact occurs naturally.

In 1994, Colombia introduced a nationwide classification system dividing the population into 6 strata based on housing characteristics and neighborhood amenities.[4] We use this exogenous cutoff as the measure of social class in our experiment: Students in strata 1 to 3 are categorized as low-SES, and those in strata 4 to 6 as high-SES (see Appendix Figure A.1 for a detailed stratum distribution of our sample).

We invite all students enrolled in two compulsory courses to participate in our experiment. Throughout the term, students meet weekly for three-hour sessions where attendance is mandatory. Both courses are university-wide graduation requirements which result in large variations in academic programs (see Appendix Table A.3) and socioeconomic backgrounds across the classrooms. This setup provides a unique opportunity

[3] Colombia has consistently ranked as one of the most unequal countries in Latin America (World Bank, 2024), with the richest decile earning 50 times more than the poorest decile (United Nations, 2023). This economic disparity is reflected by a highly stratified society with significant class inequalities and limited class mobility (Angulo et al., 2012; García et al., 2015).

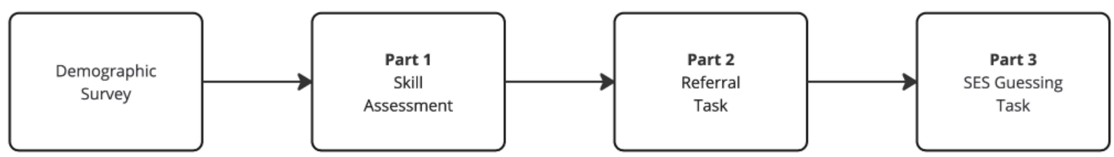
[4] Initially designed for utility subsidies from higher strata (5 and 6) to support lower strata (1 to 3), it now extends to university fees and social program eligibility. Stratum 4 neither receives subsidies nor pays extra taxes. This stratification system largely aligns with and potentially reinforces existing social class divisions (Guevara S & Shields, 2019; Uribe-Mallarino, 2008).

for collaborative inter-class contact on equal status, whose positive effects on reducing discrimination are casually documented (Lowe, 2021; Mousa, 2020; Rao, 2019).

3.3 Design

We designed an experiment to assess the peer screening ability for different skills and to measure biases related to social class. The study design consists of a single experiment with sessions organized at the classroom level (see Figure 4.2). The instructions are provided in Appendix A.2.

Figure 3.1: Experiment Timeline



Note: Participants first complete incentivized skill tests, then refer classmates for skills. In the final part, they guess the social class of their peers. This order is implemented in all sessions.

3.3.1 Skill Assessment

To understand the basis for referral decisions, we collect objective measures of cognitive and social skills. These two distinct skills are crucial for the labor market and suitable to assess given classmates interact through the term. By measuring skills before the referral stage, we eliminated the need for referred students to take additional action. Participants perform two incentivized skill tests. They have 5 minutes to complete each test. We provide test-specific instructions and an example item before participants begin. Correctly solved items increase chances to earn a fixed bonus.[5]

[5]The tests are presented in a randomized order. No performance feedback is provided. Participants see one item at a time and cannot return to previous screens once they start a test. They are not required to answer items and can skip them if they choose to do so. We elicit beliefs about performance after each test.

We use Raven's Progressive Matrices to measure cognitive skills (Raven, 1936; Raven et al., 1976). Raven's test is a well-established measure of fluid intelligence, i.e., an individual's capacity to reason and solve problems in novel situations independent of past knowledge (Schilbach et al., 2016). In this test, participants see series of images where there is a pattern with a piece that has been intentionally removed. They are tasked with choosing the piece that completes the pattern among available options. For each image, there is only one correct answer. We implement an 18-item version featuring increasingly difficult questions, with 6 response options for the first 9 items and 8 thereafter.

We measure social skills with the Multiracial Reading the Mind in the Eyes Test (MRMET) from Kim et al. (2022).[6] The test is an established measure for the ability to recognize emotions in others, and it has been previously used in economic experiments (van Leeuwen et al., 2018; Weidmann & Deming, 2021; Zárate, 2023). MRMET tends to correlate with fluid intelligence as measured by Raven's (Alan & Kubilay, 2025). It consists of photos of human faces portraying different emotions, cropped so that only the eye region is visible. Participants must choose the emotion that best describes the photo from the available answers. For each photo, there is only one correct answer and 4 response options. We administer the first 36 items in MRMET.

3.3.2 Referral Task

After the skill assessment, we create the referral task to screen for high skilled peers. For each skill, participants make incentivized referrals by nominating classmates. We first explain the measured skill accompanied by an example test item. We them provide an alphabetically ordered list of all classmates. Participants make three referral choices per skill. They are instructed to exclude themselves from referrals. A classmate may be nominated once per triad. The order in which participants refer for a skill test is randomized. We incentivize referrals with classroom-level performance rankings. The three highest-scoring classmates are designated as the top 3 for a skill. Referrers are

[6] We choose MRMET because it is a race- and gender-inclusive test suitable for application in non-WEIRD (Western, Educated, Industrial, Rich, Democratic) populations like the one we sample from. The test is based on the original RMET (Baron-Cohen et al., 2001).

eligible for a fixed bonus for referrals among the top 3.[7]

We have two between subject treatments that varies the top 3 selection. In the **Baseline** treatment, the top 3 selection is based solely on performance ranking, regardless of other participant characteristics. The **Quota** treatment modifies the top 3 selection to prioritize low-SES individuals. We reserve the first spot in the top 3 for the highest-scoring low-SES peer, and assign the remaining two places based on performance (see Table 4.1). This guarantees at least one low-SES participant in the top 3 per skill. Participants are informed about the top 3 selection mechanism before making referral choices (Appendix Figure B.1 provides illustrations explaining the treatments). Assignment to the treatment is at the individual level within each classroom. This allows comparing the effect of the treatment while keeping the referral choice set constant.

Table 3.1: Places in the Top 3 according to composition rule

	Baseline	**Quota**
Merit-only	3	2
Reserved for low-SES	0	1

3.3.3 Socioeconomic Status Guessing Task

Participants make guesses about the anticipated SES of their classmates. We inform participants that a computer algorithm randomly selects three students belonging to strata 1, 2, or 3. They are tasked with nominating the people they believe the computer could choose at random (Appendix Figure B.2 provides the illustration explaining the task). Participants select three classmates from an alphabetically ordered list containing all their classmates. This task measures the ability to distinguish SES independent of test performance, as SES identification is relevant to our study.

[7] We solve ties among the top 3 randomly. We describe only the top 3 selection mechanism and provide no feedback about the top 3 composition to participants.

3.4 Sample, Incentives, and Procedure

We invited 849 undergraduate students to participate in the experiment. Our final sample consists of 702 individuals who completed the study, resulting in an 83% participation rate.[8] We block randomized participants into treatments balancing gender and social class. Table 3.2 presents key demographic characteristics and academic performance indicators across treatments (Appendix Table A.1 illustrates the selection into the experiment). The sample is well-balanced between the **Baseline** and **Quota** conditions and we observe no statistically significant differences in any of the reported variables (all p values > 0.1). Our sample is characterized by a majority of low-SES students with about one-third of the sample being first-generation college students. The gender distribution is balanced. The mean GPA of 3.95 is consistent across both treatments.

[8] The missing students did not come to class on the day of the experiment.

Table 3.2: Balance between treatment conditions

	Baseline	**Quota**	***p***
Low-SES	59%	55%	0.297
Female	52%	47%	0.195
Cognitive score (Raven's)	10.04	10.27	0.322
Social score (MRMET)	18.45	18.50	0.886
GPA	3.95	3.95	0.828
Entry exam score	61.85	62.17	0.638
Age	19.33	19.02	0.228
First generation	34%	37%	0.386
Ethnic minority	1%	3%	0.133
Rural community	30%	27%	0.308
Scholarship	1%	1%	0.916
# semesters at university	3.18	3.17	0.916
N	368	334	702

Note: Low-SES indicates strata 1, 2, or 3. Cognitive score measures Raven's performance out of 18 questions. Social score reflects MRMET performance out of 36 questions. GPA indicates average grades out of 5. Entry exam represents the average score across reading, math, social sciences, and science components of Colombia's standardized university entrance exam ICFES. First generation indicates neither parent attended university. Rural community denotes residence in a non-urban area. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests with equal variances. All reported p-values are two-tailed.

Participants could earn bonuses worth 100,000 Pesos (about 26 US Dollars) in each part of the experiment. In the first part, we incentivized performance in the skill tests. 20% of participants were eligible for the bonus. We randomly picked one skill test for each eligible participant and drew a number between 1 and 100. The participant received the bonus if the percentage of correct answers in the selected test exceeded the drawn number. Chances of earning the bonus increased with each correctly solved question by 5.5% (=1/18) for the Cognitive Skill test and by 2.78% (=1/36) for the Social Skill test.

In the second part, we incentivized referrals among the top 3 performers. 40% of participants were eligible for the bonus. We randomly selected one skill test and one referral for each eligible participant. The participant received the bonus if their referral was among the top 3. In the third part, we incentivized the correct identification of low-SES peers. 20% of participants in each classroom were eligible for the bonus. We randomly selected one guess for each eligible participant. The participant received the bonus if their guess correctly identified a low-SES peer. Draws for the bonuses were independent meaning participants could earn multiple bonuses.

Data collection occurred during the last two weeks of April 2024. Our local partner coordinated scheduled classroom visits and recruited research assistants to administer the experiment. Students present in class on the scheduled visit dates participated. Each classroom visit constituted a separate session. There were in total 35 sessions.[9] Participants accessed the Qualtrics-based experiment using their smartphones during these visits. The median time to complete the survey was 20 minutes, with a compensation of $26 for 117 lottery winners.

3.5 Results

3.5.1 Can peers screen cognitive and social skills?

Our first goal is understanding whether higher skilled individuals get more referrals. Because every referrer nominates 3 classmates per skill, analyzing only the extensive margin, i.e., whether an individual gets a referral, is not very informative.[10] We consider the percentage share of referrals from individuals in **Baseline** condition as our dependent variable. This approach combines the intensive and extensive margins and also makes comparisons across classrooms with different sizes easier.[11]

[9] See Appendix Figures A.2a, A.2b and A.2c for the distribution of skills and GPA across classrooms and Appendix Table A.3 for diversity in program choices.

[10] Only 86 of the 849 students (10%) never get a referral for either skill.

[11] The number of participants in a classroom mechanically drives the number of total referrals that could be received by an individual. By normalizing referrals we focus on differences within classrooms.

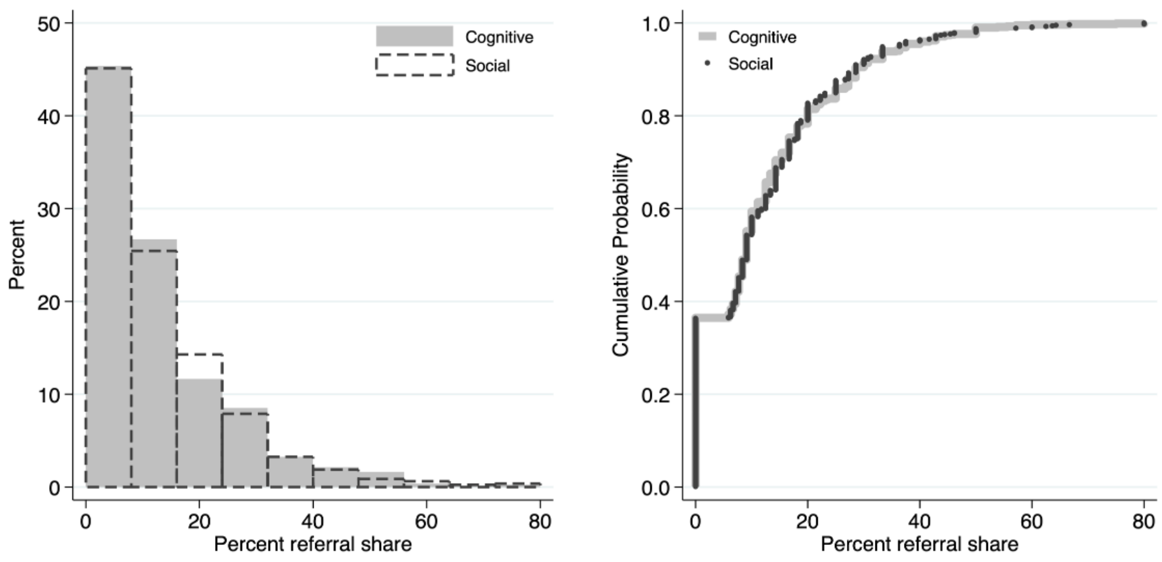
Formally, we define the percentage share of referrals received by individual i from participants j in classroom c and in **Baseline** condition ($\forall j \in B_c$) for skill s as:

$$y_{ic}^s = \frac{\sum_{j \neq i} r_{ijc}^s}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{3.1}$$

where n_c represents the number of participants in the **Baseline** condition in classroom c. The indicator r_{ijc}^s takes value 1 if participant j in the **Baseline** condition refers individual i for skill s, and 0 otherwise, and require both i and j to be in the same classroom c. The denominator $n_c - \mathbb{1}(i \in B_c)$ accounts for the maximum possible referrals that individual i could receive. If i is in the **Baseline** condition ($\mathbb{1}(i \in B_c) = 1$), we subtract one from n_c to account for the self-referral restriction.[12] This normalized measure represents the percentage of potential referrals actually received by each individual, adjusting for classroom size and treatment status. By construction, $y_i^s \in [0, 100]$ for all c, and we can compare referrals across classrooms of different sizes. Figures 3.2a and 3.2b present the distribution of our dependent variable.

[12] 33.8 percent of participants in the sample for cognitive and social skills self-referred, while explicitly instructed not to do so. In Appendix Table A.4 we compare the outcomes of those who self-refer. Self-referrers are more likely to be low-SES, and have significantly lower cognitive skill (0.2 SD) and GPA (0.25 SD). We rule out the hypothesis that self-referrers nominate themselves strategically. As self-referrerals are not informative and add noise to our estimates, we drop these instances from our paired referral-referrer sample in subsequent analyses. Self-referrering participants' remaining referral choices are kept in the dataset.

Figure 3.2: Distribution of referrals by skill in Baseline



(a) Frequency Histogram (b) ECDF

Note: Figures show the percantage of referrals recieved from participants in the **Baseline** condition for cognitive and social skills. The left panel shows the frequency histogram and the right panel shows the empirical cumulative distribution function (ECDF). A two-sample Kolmogorov-Smirnov test shows no statistically significant difference between the share of referrals received across the skill distributions ($D = 0.0363$, $p = 0.668$).

Under performance pay in the **Baseline** condition, classmates with higher scores in the skill tests should collect more referrals if classmates can screen skills. Our independent variables are the standardized skill test scores. We estimate referral percentage shares y_i^s:

$$y_i^s = \alpha^s + \beta_1^s Score_i^s + \epsilon_i^s \tag{3.2}$$

Table 3.3 illustrates our first findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the point estimates for different test scores. In column (2), a one standard deviation increase in cognitive skill score causes a 1.5 percentage point increase in the share of referrals received. On a base rate of 13%, this is a modest increase of 11.5 percent. In column (4), 95% confidence intervals rule out that a one standard deviation increase in the social skill score results in more than a 0.1 percentage point difference in the share of referrals received.

Result 1 *Participants have difficulties screening skills in the **Baseline** condition, with modest screening ability for cognitive and no screening ability at all of social skill test scores.*

Table 3.3: Share of referrals received conditional on skill test score

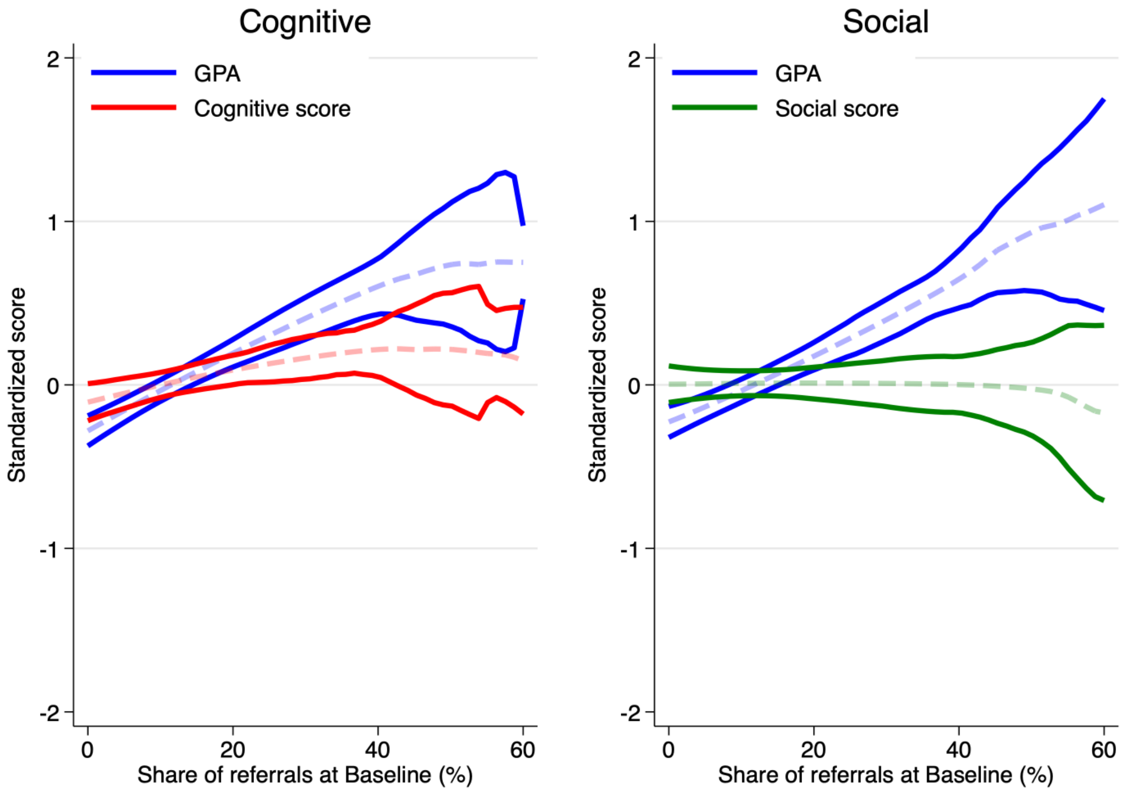
	Cognitive		Social	
	(1)	(2)	(3)	(4)
Score	1.197**	1.497***	0.037	-0.080
	(0.479)	(0.464)	(0.474)	(0.461)
Dep. var. mean	12.986	12.981	13.049	13.050
Classroom FE	No	Yes	No	Yes
R^2	0.008	0.116	0.000	0.100
Observations	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables are the percentage of referrals received relative to all referrals. "Score" refers to standardized test scores for cognitive and social skills. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

3.5.2 Grades as a proxy for skills

Absence of a clean skill-signal or the lacking the screening ability for skills may have pushed partipants to refer classmates using proxies of skills. Proxies are peer beliefs about strong correlates for skills. A potential proxy for cognitive skill (i.e., "smart students") would be the "students with good grades" in the classroom, as measured by GPA. Figure 3.3 illustrates the relationship between grades, skill score, and the share of referrals received.

Figure 3.3: Referral shares by GPA and skill test scores



Note: The left panel shows how GPA and cognitive skill scores vary with the share of cognitive skill referrals received, while the right panel shows the same for GPA and social skill score for the share of social skill referrals received. Solid lines indicate 95% confidence intervals and dashed lines indicate the means. Output is truncated at 60 percent of referral share for the sake of having meaningful confidence intervals.

The idea that grades signal cognitive skill is a common belief among researchers and practitioners alike. Yet, cognitive skill and grades are far from perfectly correlated (Heckman & Kautz, 2012; Heckman et al., 2006), and screening with such beliefs may not lead to good referrals. Indeed, GPA correlates very weakly with skill test scores in our sample (see Appendix Table A.2). We capture the screening behavior using proxies by including the standardized GPA of referrals as an independent variable. We reestimate referral percentage shares for the **Baseline** condition:

$$y_i^s = \alpha^s + \beta_1^s Skill_i^s + \beta_2^s GPA_i^s + \epsilon_i^s \tag{3.3}$$

Table 3.4 illustrates our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the difference between point estimates for skill test scores and GPA. In column (2), a one standard deviation increase in cognitive skill score causes a 1.1 percentage point increase in the share of referrals received when controlling for GPA. On a base rate of 12.8%, this is a comparable increase in magnitude of about 8.6 percent to our previous estimate in Table 3.3, and suggests cognitive skills have an independent effect on referrals. However, a one standard deviation increase in GPA causes a substantial 4.4 percentage point increase in the share of referrals received when controlling for cognitive skill score. This is an increase of four times in terms of magnitude (34 percent) when compared to cognitive skill, and suggestive of the extent to which academic performance is easier to screen among peers in our setting.

In column (4), 95% confidence intervals rule out that a one standard deviation increase in the social skill score results in more than a 0.5 percentage point difference in the share of referrals received. This is consistent with our previous estimate confirming participants cannot screen social skill scores. On the other hand, a one standard deviation increase in GPA causes a substantial 3.8 percentage point increase in the share of referrals received when controlling for social skill. This is a 30 percent increase in the share of referrals when including controls for social skill.

Result 2 *For both skills, we find strong evidence that grades act as a proxy for referral decisions.*

Table 3.4: Share of referrals received conditional on skill test score and academic performance

	Cognitive		Social	
	(1)	(2)	(3)	(4)
Score	0.873*	1.080**	-0.278	-0.527
	(0.467)	(0.455)	(0.460)	(0.409)
GPA	3.949***	4.364***	3.429***	3.789***
	(0.664)	(0.684)	(0.581)	(0.651)
Dep. var. mean	12.806	12.783	12.891	12.876
Classroom FE	No	Yes	No	Yes
R^2	0.095	0.204	0.064	0.165
Observations	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables are the percentage of referrals received relative to all referrals. "Score" refers to standardized test scores for cognitive and social skills. GPA is standardized to mean zero and unit variance. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

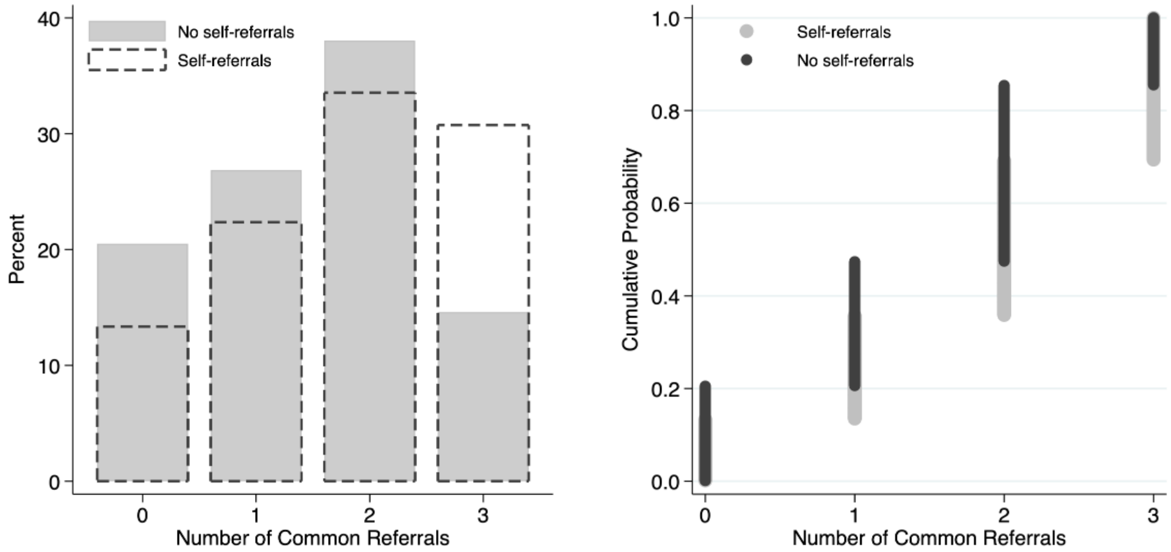
3.5.3 Types of Referrals

In this section, we expand on the diversity in referral choices to differentiate between referrers using GPA proxy and others. Despite having the opportunity to nominate up to six different classmates across two skills, referrals choices were highly concentrated. The median participant nominated two classmates in common, effectively using four of their six referral slots for the same individuals. Considering self-referrals which illustrate participants' original choices,[13] the majority of participants nominated two classmates in common for both skills, and picked themselves or someone else with almost equal probability. We visualize referral concentration by plotting the number of common referrals

[13]Self-referrals were not valid and are excluded from the main analyses.

made across skills in Figures 3.4a and 3.4b.[14]

Figure 3.4: Common referrals between skills at Baseline



(a) Frequency Histogram

(b) ECDF

Note: Figures show the distribution of common referrals with and without self-referrals. The first bar (value of 0) indicates the share of participants with 6 unique referrals. The last bar (value of 3) indicates the share of participants with 3 identical referral choices across both skills.

With such a large share common referrals across skills, it is possible that participants believed classmates with a higher score in one skill would also have a higher score in the other. Would such beliefs be accurate? There is modest ($\rho = 0.267$) correlation between the two skill test scores (see Appendix Table A.2). To understand whether making common referrals is strategic, we turn to the incentives. Participants were incentivized to pick the top 3 performers for each skill to earn a fixed bonus. Looking at the characteristics of top skilled participants in Appendix Table A.6, we find that conditional on being among the top 3 for one of the skills, only 1 in 3 participants were in the top 3 for the other skill too. This suggests *ex-post* making more than 1 common referral across skills would decrease the chances to win the bonus.

[14]In Appendix Table A.5 we compare the characteristics of referrers who make unique referrals to those who made at least one common referral. Results suggest minimal differences in GPA, skills, and social class.

A competing explanation for the amount of common referral choices between skills coupled with the notable difficulties in screening for skills would be that individuals who refer classmates twice for both skills are worse at screening. This imples the underlying heterogeneity in skill identification results in differential referral strategies where participants with a good signal for a skill choose to refer classmates only once for that skill, and those without a good signal use the grades proxy and refer classmates for both skills. We can test both hypotheses in our data: If "common" referrers -defined as those who refer an individual for both skills- are better at screening at least one of the skills, point estimates for skills in common referrals would be larger than those made uniquely for a skill. This would give credence to beliefs about correlated skills. On the other hand, if common referrers are worse in skill identification compared to unique-referrers and use GPA proxy for referrals, we can infer that they have no additional information about skills.

We compare the outcomes of participants who recieve common referrals from their classmates to those who recieve unique referrals per skill. Formally, let indicator r_{ijc}^{common} take value 1 if individual j referred individual i for both skills. The percentage share of referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) is:

$$y_{ic}^{common} = \frac{\sum_{j \neq i} r_{ijc}^{common}}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{3.4}$$

where n_c represents the number of participants in the **Baseline** condition in classroom c. The indicator r_{ijc}^{common} takes value 0 if participant j in the **Baseline** condition does not refer individual i for both skills. The denominator $n_c - \mathbb{1}(i \in B_c)$ accounts for the maximum possible "common" referrals that individual i could receive as before. Similarly, let $r_{ijc}^{s,unique}$ take value 1 if individual j referred individual i only for skill s. The percentage share of "unique" referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) for skill s is:

$$y_{ic}^{s,unique} = \frac{\sum_{j \neq i} r_{ijc}^{s,unique}}{n_c - \mathbb{1}(i \in B_c)} \times 100 \tag{3.5}$$

and it follows that for any s, percentage share of "unique" and "common" referrals received by individual i from participants in classroom c and in **Baseline** condition ($\forall j \in B_c$) must add up to the total share of referrals received:

$$y_{ic}^{s} = y_{ic}^{s,unique} + y_{ic}^{common} \tag{3.6}$$

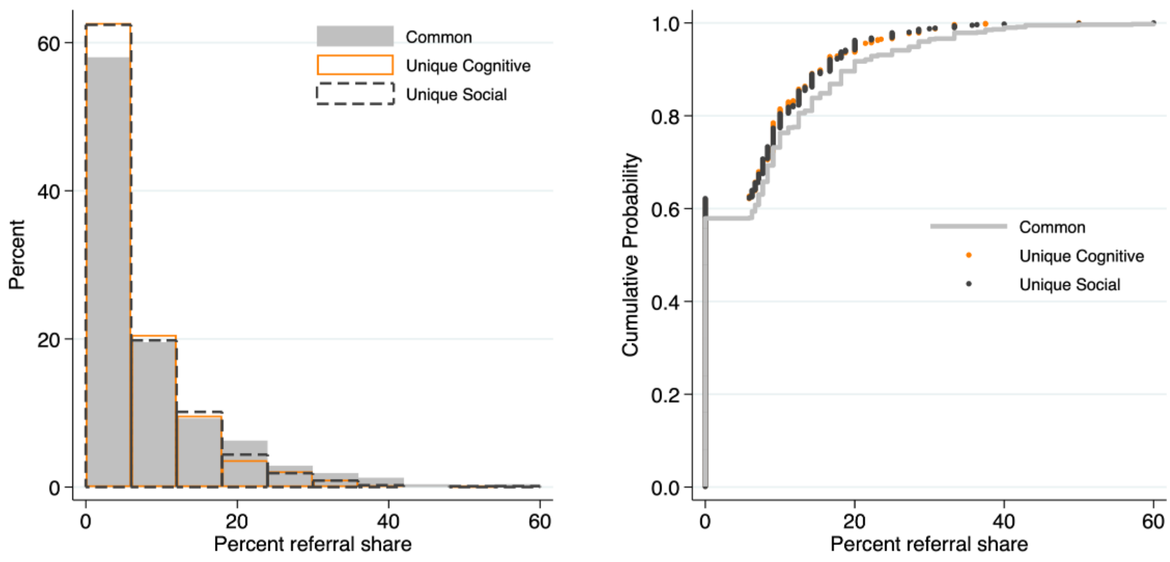
Table 3.5 shows the distribution of the referrals types in our sample. 37% of referrals fall under the type y_{ic}^{common} as pairs. This is equivalent to saying 54% of cognitive and social skill referrals were made in common. Figures 3.5a and 3.5b present the distributions of the three referral types.

Table 3.5: Distribution of Referral Types

	Frequency	Share (%)
Common	945	37.06%
Unique Cognitive	794	31.14%
Unique Social	811	31.80%
Total	2,550	100.00%

Note: Common referrals indicate the pair when the same classmate was referred for both cognitive and social skills. Unique referrals indicate when a classmate was referred for only one of the skills.

Figure 3.5: Distribution of common and unique referrals in Baseline



(a) Frequency Histogram

(b) ECDF

Note: Figures show the percantage of referrals recieved from participants in the **Baseline** condition depending on the referral type. The left panel shows the frequency histogram and the right panel shows the empirical cumulative distribution function (ECDF). Two-sample Kolmogorov-Smirnov tests show no statistically significant differences between the share of referrals received between "unique" cognitive and "unique" social referral distributions ($D = 0.0125$, $p = 1.000$) as well as "common" referrals ($D = 0.0602$, $p = 0.111$ for cognitive and $D = 0.0551$, $p = 0.177$ for social).

We regress Equation 3.3 for our three new dependent variables and report our findings in Table 3.6. Our preferred specification includes classroom fixed effects. The comparison of interest is the skill test scores and GPA estimates across columns. In column (2), we find that a one standard deviation increase in GPA causes a 3.7 percentage point increase in the share of "common" referrals received when controlling for skill test scores. This is a substantial 50 percent increase on a base rate of 7.4%. Cognitive skills remain statistically insignificant and social skills show a marginally significant negative coefficient, suggesting that participants who nominate the same individuals for both skills primarily make referrals based on academic performance.

For participants who receive unique cognitive skill referrals, in column (4), we find that a one standard deviation increase in GPA causes a 0.75 percentage point increase in the share of referrals when controlling for cognitive skill test score. A one standard deviation increase in cognitive skill test score causes a larger 1.1 percentage point increase in referrals when controlling for GPA. These are respectively 14 and 20 percent increases in the share of referrals received, and suggest participants are able to screen higher skilled peers when uniquely referring for cognitive skill. The lower base rate of 5.4% compared to 7.4% in column (2) suggests less than half of referrals came from "unique" referrals. The GPA estimate is five times smaller in magnitude compared to column (2), and suggests a smaller weight put on the grades proxy. Nevertheless, the comparable magnitudes of GPA and cognitive skill point estimates still suggest participants refer peers with higher grades much more often than the correlation between the two supported by the data ($\rho = 0.085$). There is heterogeneity in skill identification ability when uniquely referring for cognitive skill.

For participants who receive unique social skill referrals, in column (6), 95% confidence intervals rule out that a one standard deviation increase in social skill test score or GPA result in more than a 0.1 percentage point difference in the share of referrals received. These results further support our previous finding that peers cannot screen social skills in our sample, and do not attempt to screen social skills with the GPA proxy.

Result 3 *The majority of participants nominate the same individuals in common for both skills, cannot screen for skills and refer instead using the GPA proxy.*

Result 4 *Those who refer uniqely for cognitive skill can identify the skill test score, and drive the entierity of the results in terms of peer skill identification. Still, they confound cognitive skill with academic performance, and put comparable weights on the two. Those who refer uniqely for social skill can neither screen social skill or use the GPA proxy.*

Table 3.6: Share of "common" versus "unique" referrals received conditional on skill test score and academic performance

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
GPA	3.172***	3.670***	0.801**	0.752*	0.260	0.108
	(0.464)	(0.501)	(0.391)	(0.401)	(0.334)	(0.360)
Cognitive score	-0.042	0.139	1.006***	1.084***		
	(0.416)	(0.388)	(0.270)	(0.281)		
Social score	-0.353	-0.553*			0.086	-0.011
	(0.304)	(0.272)			(0.381)	(0.357)
Dep. var. mean	7.407	7.382	5.400	5.401	5.485	5.493
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.093	0.194	0.028	0.130	0.001	0.090
Observations	665	665	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Independent variables are the respective standardized test scores for skills and GPA. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

3.5.4 Social class bias across common and unique referral types

In this section, we analyze referrals from the perspective of social class while accounting for the referral types described in this part. Based on the referral types from the previous section, we document the existence of a social class bias in referrals when controlling for skill test scores and academic performance at **Baseline**. Our dependent variables are the percentage shares of referrals received at **Baseline** as defined in Equation 3.6, and we include a social class dummy for the participant receiving the referrals. We estimate for our three dependent variables:

$$y_i^s = \alpha^s + \beta_1^s GPA_i + \beta_2^s Score_i^s + \beta_3^s SES_i + \epsilon_i^s \tag{3.7}$$

Table 3.7 summarizes our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the SES estimates for the three referral strategies. In column (2), controlling for skill test scores and GPA, the point estimate for low-SES is not statistically significant. Skill score and GPA estimates are robust to the inclusion of this variable and remain close to those in Table 3.6.

For participants who receive unique cognitive skill referrals in column (4), we find that being low-SES causes a 1.8 percentage point decrease in the share of referrals when controlling for cognitive skill and GPA. This is a substantial 28 percent difference in the share of referrals received, confirming participants are biased against low-SES peers when uniquely referring for cognitive skill. Skill test scores and GPA estimates are robust to the inclusion of this variable. GPA and low-SES are not confounders as there are no significant differences across social classes in terms of GPA (see Appendix Figure A.3). The low-SES bias is consistent with the data where low-SES students underperform in the cognitive skill test (see Appendix Figure A.4a).

For participants who receive unique social skill referrals, in column (6), the point estimate for low-SES is not statistically significant. GPA and social skill estimates remain similar to those in Table 3.6. The finding that low-SES students underperform across skill dimensions is also consistent with earlier research (Falk et al., 2021), though we find that low-SES bias manifests only in unique cognitive skill referrals.

Result 5 *We document a sizeable low-SES bias for unique cognitive skill referrals when controlling for cognitive skill test score and academic performance of peers.*

Table 3.7: Share of "common" versus "unique" referrals received conditional on skill test score, academic performance, and social class

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
GPA	3.170***	3.663***	0.797**	0.766*	0.260	0.111
	(0.462)	(0.499)	(0.386)	(0.388)	(0.334)	(0.360)
Cognitive score	0.000	0.167	0.869***	0.973***		
	(0.411)	(0.382)	(0.261)	(0.274)		
Social score	-0.306	-0.524*			0.047	-0.027
	(0.315)	(0.286)			(0.372)	(0.354)
Low-SES	0.799	0.568	-2.017***	-1.814**	-0.549	-0.260
	(0.939)	(0.934)	(0.711)	(0.713)	(0.610)	(0.593)
Dep. var. mean	6.948	7.056	6.558	6.442	5.800	5.642
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.094	0.194	0.044	0.142	0.002	0.090
Observations	665	665	665	665	665	665

Note: Classroom-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Independent variables are the respective standardized test scores for skills, GPA, and a dummy for low socioeconomic status. Sample restricted to 665 individuals for whom we have complete administrative and experimental data.

3.5.5 Social class bias and the Quota treatment

In the following empirical specification, we document whether there is a social class bias in aggregate, and whether the **Quota** treatment causes referral shares of low-SES participants to change when controlling for skills and academic performance. We hypothesized that the **Quota** treatment should increase referrals to low-SES peers because of the additional incentive to refer low-SES. The dependent variable is the percentage share of referrals received as defined for the **Baseline** treatment in Equation 3.1, now extended to the referrals from the **Quota** treatment. It is trivial to see y^s_{ic} can also be calculated for the **Quota** treatment as participants in every classroom are randomized into either treatment. Now, every participant is observed twice in the data for the share of referrals they received from participants in either treatment. We add a treatment dummy to indicate whether the referrals came from participants in the **Baseline** or the **Quota** treatment. We also add a social class dummy for the participant receiving the referrals to our specification and estimate:

$$y^s_i = \alpha^s + \beta^s_1 Quota_i + \beta^s_2 SES_i + \beta^s_3 (Quota_i \times SES_i) + \beta^s_4 Score^s_i + \beta^s_5 GPA_i + \epsilon^s_i \quad (3.8)$$

Table 3.8 illustrates our findings. Our preferred specification includes classroom fixed effects. Our comparison of interest is the effect of the **Quota** treatment on low-SES peers. In column (2) for cognitive skill, we find that being low-SES decreases the share of referrals received by about 1.3 percentage points when controlling for the skill test score and academic performance. This difference is not statistically significant, but its direction and magnitude suggests a relatively large bias against low-SES classmates: A one standard deviation increase in cognitive skill test score has a similar magnitude (0.8 percentage points). This finding suggests the low-SES bias is driven by those who made unique cognitive referrals but it is not large enough to carry over to all cognitive skill referrals considered together. In column (4) for social skill, we find that being low-SES has no statistically significant effect on the share of referrals received when controlling for the skill test score and academic performance.

Result 6 *The low-SES bias is not large enough to carry over to all cognitive skill referrals when referrals are aggregated.*

Table 3.8: Share of referrals received by treatment, controlling for skill test score, academic performance, and social class

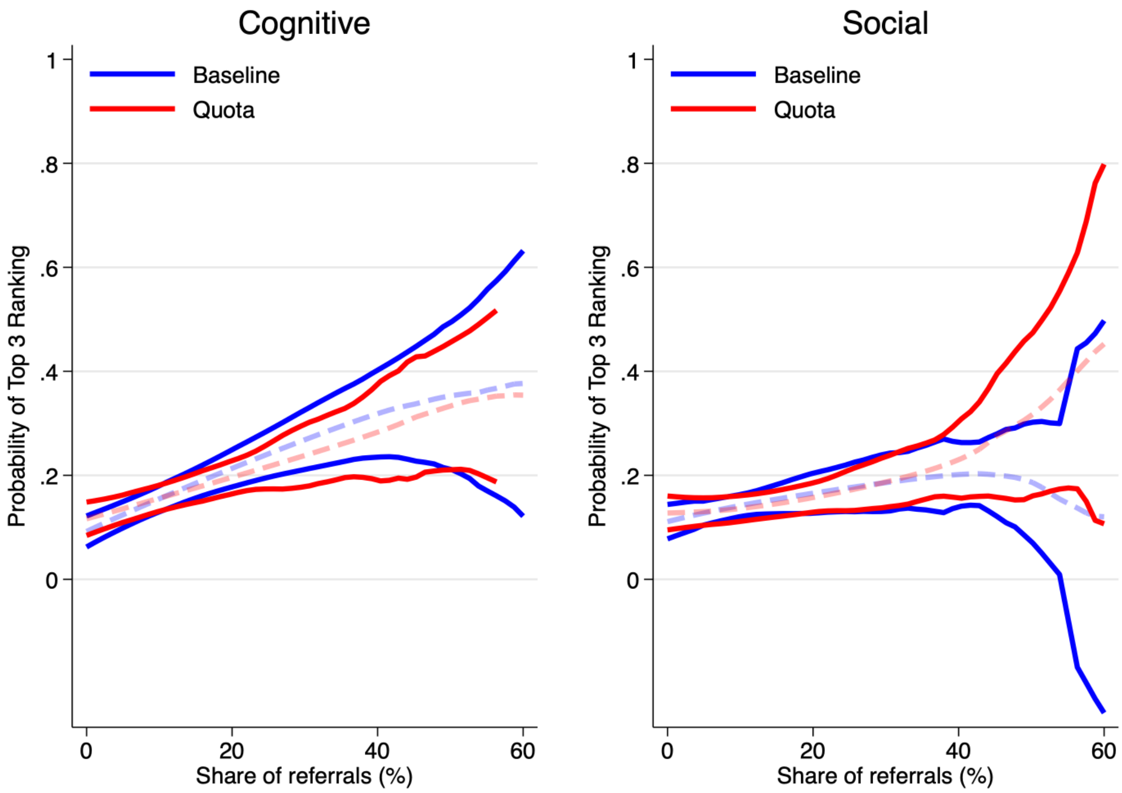
	Cognitive		Social	
	(1)	(2)	(3)	(4)
Quota	-0.073	-0.073	0.299	0.299
	(0.755)	(0.755)	(0.716)	(0.716)
Low-SES	-1.230	-1.276	0.364	0.324
	(1.079)	(1.014)	(1.282)	(1.361)
Quota $\times$ Low-SES	-0.167	-0.167	-0.835	-0.835
	(1.117)	(1.117)	(1.181)	(1.181)
Score	0.594	0.811*	0.201	-0.006
	(0.448)	(0.424)	(0.426)	(0.458)
GPA	3.184***	3.522***	2.819***	3.174***
	(0.517)	(0.552)	(0.493)	(0.621)
Dep. var. mean	13.551	13.558	12.706	12.714
Classroom FE	No	Yes	No	Yes
R^2	0.060	0.158	0.044	0.134
Observations	1,330	1,330	1,330	1,330

Note: Standard errors in parentheses are clustered at both classroom and individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables are the percentage share of referrals received for cognitive and social skills. Quota is a dummy for the referrals received from classmates in the **Quota** treatment. Low-SES is a dummy for participant's socioeconomic status. Remaining independent variables are the respective standardized test scores for skills and GPA. Sample includes 1,330 observations with complete administrative and experimental data.

3.5.6 Quota treatment and referral productivity

As any intervention that changes the nomination decisions in terms of SES composition should not reduce the productivity of referrals, the equity-efficiency traedeoff is a valid concern for the **Quota** treatment. To address it, in Figure 3.6, we plot the share of referrals received across the two conditions and the probability of being among the Top 3 in the classroom for either skill. We find first that the slopes of the distributions are always positive for the **Quota** treatment, indicating a positive relationship with the share of referrals received. Second, a two-sample Kolmogorov-Smirnov test reveals no statistically significant differences in the distribution of referrals between the two conditions for both cognitive and social referrals. These suggest that the **Quota** treatment does not impact the positive relationship between the share referrals received and productivity in skills.

Figure 3.6: Referral shares and the probability of being in the Top 3



Note: The left panel shows how Baseline and Quota referral shares vary with the probability of being in the Top 3 of the classroom for cognitive skill scores, while the right panel shows the same figure for social skill scores. Solid lines indicate the 95% confidence intervals, with dashed lines representing means. The output is truncated at 60 percent of referral share to ensure meaningful confidence intervals. Two-sample Kolmogorov-Smirnov tests reveal no statistically significant differences in the distribution of referrals between Baseline and Quota conditions for both cognitive referrals ($D = 0.0351$, $p = 0.710$) and social referrals ($D = 0.0439$, $p = 0.427$).

3.5.7 Effects of the Quota treatment across referral types

Effects of the social class bias gets diluted across common and unique referral types. A large proportion of participants -"common" referrers- who struggle with skill identification and screen for skills using the academic performance proxy. But there are no SES differences for GPA in our sample. When referrals are made with academic performance in mind, it seems reasonable not to observe a negative bias against low-SES. Then what about skills, knowing that high-SES score higher in both measures?

We observe a bias in undersampling from equally well performing low-SES only for "unique" cognitive skill referrals, where referrers screen better compared to "unique" social skill referrals. We expect the **Quota** treatment be effective in increasing referrals for low-SES only in a scenario where the skill can be screened, and turn toward our classification of different referral types to test this hypothesis. To get clearer estimates for the effects of the **Quota** treatment on low-SES referrals, we re-estimate the shares of "common" and "unique" referrals. Following the same logic in the section before, we observe every participant twice in each specification, and add a treatment dummy to indicate whether the referrals came from referrers in the **Baseline** or the **Quota** treatment. We keep the social class dummy and regress Equation 3.8 for the three dependent variables.

Table 3.9 illustrates our findings. Our preferred specification includes classroom fixed effects. The comparison of interest is the SES of the participant receiving the referrals and the effect of the **Quota** treatment across "common" and "unique" referral types. In column (2), for participants who refer the same peers in common using the academic performance proxy, we find no statistically significant effect of participant SES or the **Quota** on the referrals share when controlling for skill test scores and academic performance.

For unique cognitive skill referrals, in column (4), we find that being low-SES in the **Baseline** treatment reduces the percentage share of referrals received by 1.9 percentage points when controlling for the skill test score and academic performance. This is a very large effect size which translates to a decrease in referral share by 29 percent on a base rate of 6.5%, and is similar to the one found in Table 3.7. In turn, the **Quota** treatment increases referrals to low-SES by 1.42 percentage points when controlling for the skill test score and academic performance. This is also a large effect size that results in an increase in low-SES referral share by 22 percent.

For participants who make unique social skill referrals, in column (6), we find no statistically significant effect of participant SES or the **Quota** on the referrals share when controlling for the skill test score and academic performance. These are in accordance with our previous findings that social skills cannot be identified in our setting and it is possible that we do not observe the low-SES bias in this skill domain for this reason.

Result 7 *There is a bias against low-SES peers only for the skill that is well-identified by peers, and in which low-SES underperform. We find no evidence of a bias when referrals are made based on academic performance where both social classes perform equally well.*

Result 8 *The bias in unique cognitive skill referrals is partially alleviated by the* ***Quota*** *treatment. Because there is remarkable heterogeneity in the ability to detect SES for both social classes (see Appendix Figure A.5), this significant increase in low-SES referrals is satisfying in our setting.*

Table 3.9: Share of "common" and "unique" referrals received by treatment, controlling for skill test score, academic performance, and social class

	Common		Unique Cognitive		Unique Social	
	(1)	(2)	(3)	(4)	(5)	(6)
Quota	0.436	0.436	-0.509	-0.509	-0.136	-0.136
	(0.817)	(0.817)	(0.598)	(0.598)	(0.523)	(0.523)
Low-SES	0.857	0.598	-2.074***	-1.891**	-0.510	-0.256
	(0.920)	(0.897)	(0.722)	(0.710)	(0.613)	(0.594)
Quota × Low-SES	-1.584	-1.584	1.417**	1.417**	0.750	0.750
	(1.159)	(1.159)	(0.656)	(0.656)	(0.717)	(0.717)
Cognitive score	-0.079	0.095	0.658***	0.739***		
	(0.374)	(0.346)	(0.201)	(0.210)		
Social score	0.062	-0.091			0.158	0.061
	(0.283)	(0.236)			(0.312)	(0.269)
GPA	2.322***	2.727***	0.858***	0.804**	0.502*	0.439
	(0.330)	(0.366)	(0.312)	(0.340)	(0.278)	(0.292)
Dep. var. mean	6.952	7.080	6.591	6.488	5.765	5.623
Classroom FE	No	Yes	No	Yes	No	Yes
R^2	0.052	0.139	0.028	0.099	0.005	0.071
Observations	1,330	1,330	1,330	1,330	1,330	1,330

Note: Standard errors in parentheses are clustered at both classroom and individual level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable in columns (1)-(2) is the percentage share of "common" referrals received from the same referrer, in columns (3)-(4) "unique" referral share for cognitive skill, and in columns (5)-(6) for social skill. Quota is a dummy for the referrals received from classmates in the **Quota** treatment. Low-SES is a dummy for participant's socioeconomic status. Remaining independent variables are the respective standardized test scores for skills and GPA. Sample includes 1,330 observations with complete administrative and experimental data.

3.6 Conclusion

In this paper, we study how accurately individuals assess productivity of their peers across different skill dimensions and whether these assessments systematically disadvantage low-SES individuals in a diverse university setting. Through a lab-in-the-field experiment that isolates screening ability, we find that the accuracy of peer productivity assessments varies significantly across skill types, with implications for referral-based screening.

Our findings reveal that peers can effectively identify cognitive skills but struggle to assess social skills in their classmates. This differential screening ability appears to stem from the inherent challenges in evaluating interpersonal capabilities compared to cognitive abilities. When faced with uncertainty in skill assessment, peers often rely on observable proxies like academic performance which may be misleading. This suggests that the effectiveness of peer assessments depends crucially on how discernible the target skill is, rather than indicating a fundamental limitation of referrals as a screening mechanism.

These results complement the broader literature showing referrals' effectiveness in worker screening by highlighting how skill visibility affects assessment accuracy. While previous work demonstrates that referrals successfully identify productive workers overall (Pallais & Sands, 2016), our findings suggest their effectiveness may vary across different dimensions of human capital. This variation is particularly relevant given the growing importance of social skills in the labor market as found in other research (Deming, 2017). Our evidence also supports earlier evidence that accurate assessment of social skills remains challenging (Bassi & Nansamba, 2022), suggesting the need for either longer periods of interaction to discern these skills or development of alternative assessment methods that can better capture interpersonal capabilities in referral settings.

Looking forward, our findings suggest several implications for improving screening mechanisms in similar settings. First, institutions that implement referral programs may need to develop complementary tools for evaluating less visible skills like interpersonal capabilities, perhaps in the likes of the social skill certificates in Bassi and Nansamba (2022). Second, our results on social class bias - finding it only in unique cognitive skill referrals and its mitigation through quota incentives without compirimising productivity- indicate that targeted interventions can effectively address specific biases without compromising the overall screening process. Future research could investigate how to optimize referral programs to leverage their strengths in identifying easier to discern skills while developing better methods for assessing harder to observe skills.

References

Alan, S., & Kubilay, E. (2025). Empowering adolescents to transform schools: Lessons from a behavioral targeting. *American Economic Review*, *115*(2), forthcoming. doi: 10.1257/aer.20240374

Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en colombia. *Documentos CEDE*.

Balafoutas, L., & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, *335*(6068), 579–582.

Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, *77*(4), 1047–1094.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251. (Publisher: Cambridge University Press) doi: 10.1017/S0021963001006643

Bassi, V., & Nansamba, A. (2022). Screening and signalling non-cognitive skills: experimental evidence from uganda. *The Economic Journal*, *132*(642), 471–511.

Beaman, L., Keleher, N., & Magruder, J. (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics*, *36*(1), 121–157. doi: 10.1086/693869

Beaman, L., & Magruder, J. (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, *102*(7), 3574–3593. doi: 10.1257/aer.102.7.3574

Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2019). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, *86*(1), 191–239.

Beugnot, J., & Peterlé, E. (2020). Gender bias in job referrals: An experimental test. *Journal of Economic Psychology*, *76*, 102209. doi: 10.1016/j.joep.2019.102209

Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American economic review*, *94*(3), 426–454.

Calvó-Armengol, A., & Jackson, M. O. (2007). Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, *132*(1), 27–46.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... others (2022a). Social capital II: determinants of economic connectedness. *Nature*, *608*(7921), 122–134.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... others (2022b). Social capital I: measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The quarterly journal of economics*, *132*(4), 1593–1640.

Deming, D. J. (2023). Multidimensional human capital and the wage structure. *National Bureau of Economic Research*.

Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., & Deckers, T. (2021). Socioeconomic status and inequalities in children's iq and economic preferences. *Journal of Political Economy*, *129*(9), 2504–2545.

García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la cuna: La movilidad social a través de la educación en los municipios de colombia. *Documentos CEDE*.

Guevara S, J. D., & Shields, R. (2019). Spatializing stratification: Bogotá. *Ardeth. A Magazine on the Power of the Project*(4), 223–236.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, *19*(4), 451–464.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, *24*(3), 411–482.

Hederos, K., Sandberg, A., Kvissberg, L., & Polano, E. (2024). Gender homophily in job referrals: Evidence from a field study among university students. *Labour Economics*, *87*, 102461. doi: 10.1016/j.labeco.2024.102461

Hudson, R. A., & Library of Congress (Eds.). (2010). *Colombia: a country study* (5th ed.). Washington, D.C: Federal Research Division, Library of Congress: For sale by the Supt. of Docs., U.S. G.P.O. Retrieved from the Library of Congress, https://www.loc.gov/item/2010009203/.

Kim, H., Kaduthodil, J., Strong, R. W., Germine, L., Cohan, S., & Wilmer, J. B. (2022). *Multiracial Reading the Mind in the Eyes Test (MRMET): an inclusive version of* *an influential measure* (preprint). doi: 10.31219/osf.io/y8djm

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–128.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, *111*(6), 1807–1844.

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in america: Changes in core discussion networks over two decades. *American sociological review*, *71*(3), 353–375.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, *27*(1), 415–444.

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*.

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. *Science*, *369*(6505), 866–870.

Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? affirmative action in light of gender differences in competitiveness. *Management Science*, *59*(1), 1–16.

Pallais, A., & Sands, E. G. (2016). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, *124*(6), 1793–1828. doi: 10.1086/688850

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, *109*(3), 774–809.

Raven, J. C. (1936). *The Performances of Related Individuals in Tests Mainly Educative and Mainly Reproductive Mental Tests Used in Genetic Studies* (PhD Thesis). University of London (King's College).

Raven, J. C., Raven, J., & Court, J. H. (1976). *Standard progressive matrices: sets A, B, C, D & E.* Oxford, UK: Oxford Psychologists Press.

Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The Psychological Lives of the Poor. *American Economic Review*, *106*(5), 435–440. doi: 10.1257/aer.p20161101

Stansbury, A., & Rodriguez, K. (2024). The class gap in career progression: Evidence from US academia. *Working Paper*.

United Nations. (2023). *Social panorama of latin america and the caribbean 2023: labour inclusion as a key axis of inclusive social development.* ECLAC and United Nations. Retrieved from https://www.cepal.org/es/publicaciones/68702-panorama-social-america-latina-caribe-2023-la-inclusion-laboral-como-eje-central

Uribe-Mallarino, C. (2008). Estratificación social en bogotá: de la política pública a la dinámica de la segregación social. *Universitas humanistica*(65), 139–172.

van Leeuwen, B., Noussair, C. N., Offerman, T., Suetens, S., van Veelen, M., & van de Ven, J. (2018). Predictably Angry—Facial Cues Provide a Credible Signal of Destructive Behavior. *Management Science*, *64*(7), 3352–3364. doi: 10.1287/mnsc.2017.2727

Weidmann, B., & Deming, D. J. (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica*, *89*(6), 2637–2657. doi: 10.3982/ECTA18461

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, *96*(5), 849–861.

Witte, M. (2021). Why do workers make job referrals? experimental evidence from ethiopia. *Working Paper*.

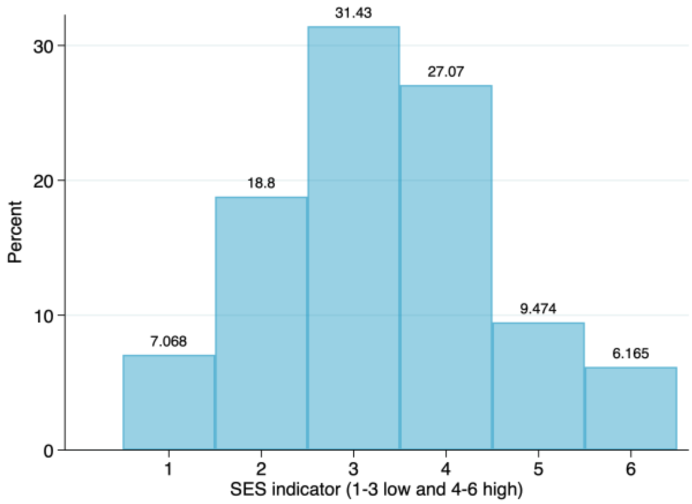
World Bank. (2024). *Regional poverty and inequality update spring 2024* (Poverty and Equity Global Practice Brief). Washington, D.C.: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269

Zárate, R. A. (2023, July). Uncovering peer effects in social and academic skills. *American Economic Journal: Applied Economics*, *15*(3), 35–79. doi: 10.1257/app.20210583

A.1 Additional Figures and Tables

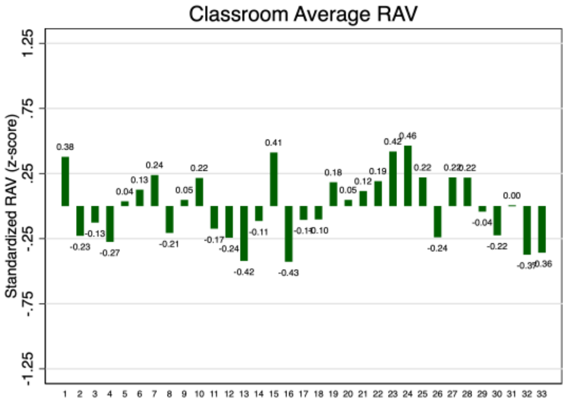
A.1.1 Additional Figures

Figure A.1: Stratum distribution of the sample

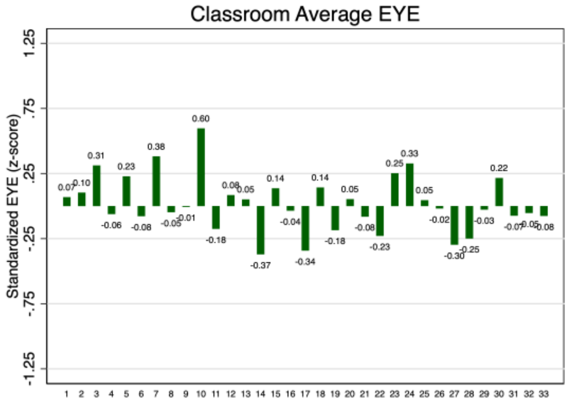


Note: This figure shows the distribution of strata in the sample of students that participated in the study.

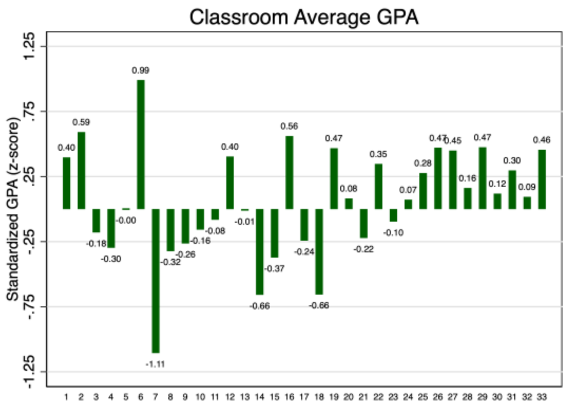
(a) Cognitive score across classrooms



(b) Social score across classrooms

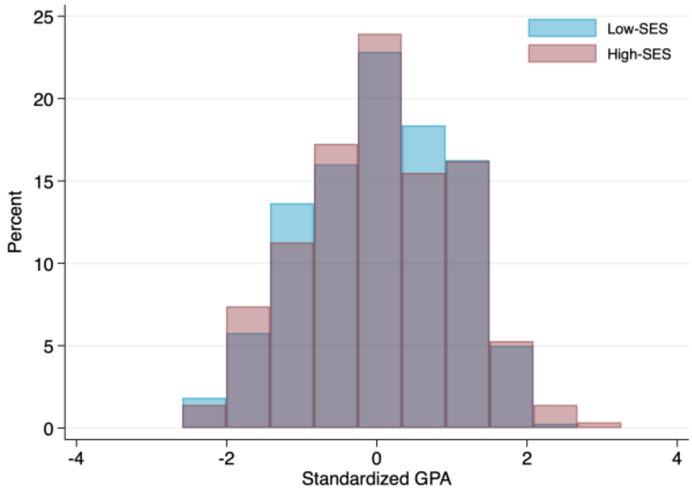


(c) GPA across classrooms



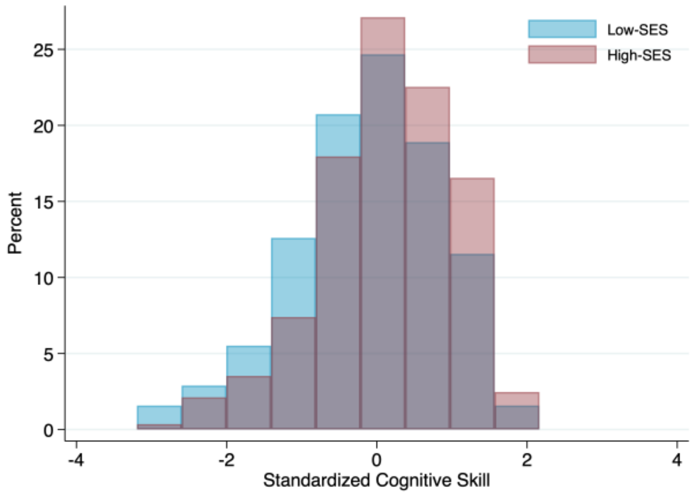
Note: These figures show the respective distribution of standardized scores for cognitive skill, social skill, and GPA across sampled classrooms.

Figure A.3: GPA by SES

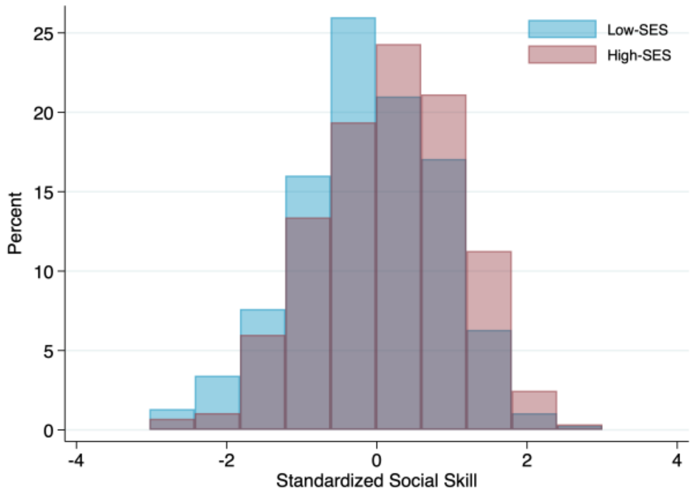


Note: This figure shows the distribution of GPA across SES. There are no significant differences in the mean standardized GPA scores between high-SES and low-SES participants (t test $p = 0.695$).

(a) Cognitive score by SES

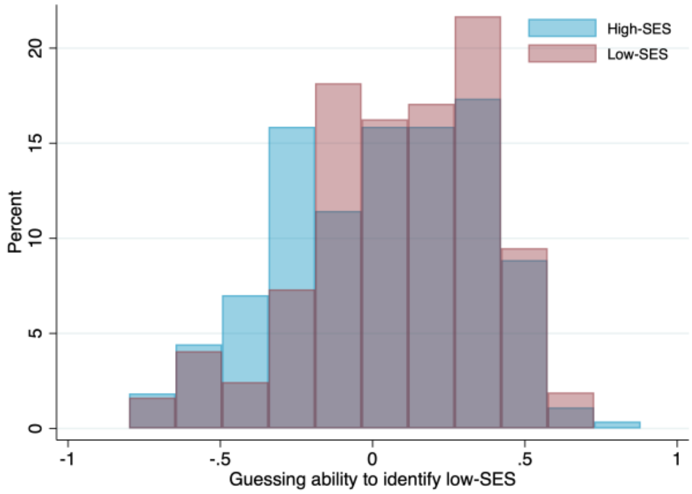


(b) Social score by SES



Note: These figures show the respective distribution of cognitive and social skills across SES. High social class outperform Low-SES in both skills (t tests have p values < 0.001). We can visually verify that larger share of high-SES in quantiles above median for both skills.

Figure A.5: Distribution of guessing ability across SES



Note: This figure shows the distribution of the guessing ability across SES. We calculate the guessing ability as the share of succesful low-SES guesses minus the expected probability of randomly drawing low-SES in class c. A score of 0 indicates an accuracy as good as random draws, below 0 drawing worse than chance, and above 0 better than chance. There are significant differences in the mean guessing ability between high-SES (M = 0.022, SD = 0.325, n = 271) and low-SES participants (M = 0.093, SD = 0.302, n = 369), $t(638) = -2.85$, $p = 0.005$, $d = 0.226$. Low-SES participants have higher guessing ability compared to their high-SES counterparts, with a mean difference of 7 percentage points.

A.1.2 Additional Tables

Table A.1: Selection into the experiment

	Sample	**Missing**	***p***
Referral share (both skills)	0.127	0.043	0.000
GPA (standardized)	0.044	-0.273	0.001
Entry Exam (standardized)	0.028	-0.168	0.046
# Semesters at UNAB	3.171	3.188	0.884
Age	19.182	20.287	0.001
Female	49.8%	48.5%	0.788
Ethnic Minority	2.1%	4.4%	0.114
Rural Community	28.8%	31.6%	0.501
Has Scholarship	0.8%	0.7%	0.899

Note: Values for female, ethnic minority, rural community, and scholarship represent percentage proportions. All other variables represent means. p-values for gender, ethnic, rural, and scholarship are from two-sample tests of proportions. For all other variables, p-values are from two-sample t-tests with equal variances. All tests compare the sample and missing students. All reported p-values are two-tailed.

Table A.2: Correlation between GPA, entry exam, and skill test scores

	GPA	Cognitive score	Social score	Entry Exam
GPA	1.000			
Cognitive score	0.083	1.000		
Social score	0.091	0.266	1.000	
Entry Exam	0.229	0.403	0.267	1.000

Note: Pairwise correlation between GPA, entry exam, and skill test scores. Sample is restricted to 655 participants with complete administrative and experimental data.

Table A.3: Between-Classroom Variation in Academic Programs

Statistic	Most common program share
Mean	0.424
Standard Deviation	0.216
10th percentile	0.174
25th percentile	0.292
Median	0.345
75th percentile	0.533
90th percentile	0.696
# classrooms with share 1	3
Most diverse classroom	0.154
# classrooms	35

Note: Table shows the distribution of academic programs across classrooms, measured by the share of students from the most common program in each classroom. Three classrooms are completely homogeneous (share = 1). In the median classroom, the most common program accounts for 34.5% of students. The most diverse classroom has only 15.4% of students in the same program. Data based on 849 students across 35 classrooms.

Table A.4: Characteristics of self-referrers

	No self-referral	Any self-referral	Δ	p
GPA	0.132	-0.120	0.252	0.002
	(1.003)	(0.966)		
Cognitive score	0.087	-0.118	0.205	0.013
	(0.988)	(1.023)		
Social score	0.034	-0.038	0.072	0.374
	(1.003)	(0.959)		
Low-SES	0.605	0.511	0.094	0.021
	(0.490)	(0.501)		
N	440	225	665	
Share (%)	66.2	33.8	100	

Note: Table compares standardized scores between participants who self-referred at least once (N = 225) and those who did not (N = 440). Positive differences indicate higher scores for those who never self-referrered. p-values from two-sided t-tests (GPA, Cognitive Skill, Social Skill) and proportion test (Low-SES). The results suggest self-referrers have significantly lower cognitive skills and GPA, and are more likely to be low-SES. Standard deviations in parentheses, samples restricted to participants with complete administrative and experimental data.

Table A.5: Characteristics of participants who make overlapping referrals

	Unique referrals	Common referrals	Δ	p
GPA	0.057	0.045	0.012	0.903
	(0.983)	(1.009)		
Cognitive score	0.110	0.024	0.086	0.371
	(1.005)	(0.979)		
Social score	-0.014	0.033	-0.047	0.621
	(0.938)	(0.981)		
Low-SES	0.530	0.597	-0.067	0.164
	(0.501)	(0.491)		
N	132	512	644	
Share (%)	20.5	79.5	100	

Note: Table compares characteristics between participants who made at least one overlapping referral (N = 512) to those who did not (N = 132, 20.5%). Overlapping referrals indicate cases where a participant referred the same classmate once for cognitive or social skills. Positive differences indicate higher scores for those who made no overlapping referrals. The results suggest minimal differences across all variables. p-values from two-sided t-tests (GPA, Cognitive Skill, Social Skill) and proportion test (Low-SES). Standard deviations in parentheses, sample restricted to participants with complete administrative and experimental data.

Table A.6: Characteristics of Top Performers and Referrals

	Cognitive		Social		Both
	Top 3	Referrals	Top 3	Referrals	Top 3
Cognitive score	1.223	0.112	0.383	0.058	1.201
	(0.419)	(1.009)	(0.922)	(1.015)	(0.458)
Social score	0.357	0.086	1.340	0.042	1.391
	(0.923)	(0.996)	(0.395)	(1.009)	(0.453)
GPA	0.277	0.251	0.264	0.212	0.551
	(0.990)	(1.021)	(1.046)	(1.004)	(0.897)
Low-SES	0.457	0.532	0.456	0.555	0.500
	(0.500)	(0.499)	(0.500)	(0.497)	(0.507)
N	129	1,759	114	1,775	36
Share (%)	20.0	100	17.7	100	5.6

Note: Table shows characteristics of students ranked in the top 3 of their classroom and average characteristics of referred students, by skill. Standard deviations in parentheses. Sample restricted to participants with complete administrative and experimental data. All continous variables are standardized.

A.2 Experiment

We include the English version of the instructions used in Qualtrics. Participants saw the Spanish version. Horizontal lines indicate page breaks, and clarifying comments are inside brackets.

Please enter the password:

[classroom-specific password sent to each participant the day before data collection]

Welcome

Welcome to this study organized by the Social Bee Lab. You have been invited to participate in a survey where you can make a series of decisions. The study takes approximately 20 minutes to complete. During the study, you should not communicate with any other students. If you have any questions at any time, please raise your hand. One of the assistants will help you privately.

In this study, you can win bonus money depending on your choices. In total, we will draw [classroom-specific number equal to 40% of class size] bonuses of 100.000 pesos among the participants of this classroom. It is also possible for the same person to win more than one voucher. The following screens will detail how the bonus draw will be conducted. The UNAB finance office will make the payment of the vouchers through Nequi.

All your decisions in this survey will be anonymized. Therefore, the answers you provide will not affect your grades in this class or your records at the university. We will use your personal information to determine the bonus allocation, but after that, we will remove any data that identifies you.

This survey has several parts. Each of these parts has specific instructions. Please read the instructions for each part carefully because they describe how you can earn bonuses. This study has been approved by the [omitted for anonymous review] on the condition that all the information we provide is true and all the bonuses we offer are real.
On the next screen, we present you with an informed consent form that you must accept to participate in this study.

Informed Consent

You have been invited to participate in a study to learn more about how people make decisions in common scenarios.

This study is conducted by [omitted for anonymous review] and the Social Bee Lab at UNAB. The purpose of this study is to broaden our understanding of how people make decisions.

Participation in this study is voluntary. You may opt-out at any time. No known risks are associated with your participation in this project beyond those of everyday life. Apart from the monetary bonuses that will be drawn, participation has no direct benefits.

The Social Bee Lab is in charge of data collection. Your answers in this study are anonymous and will not be shared with anyone. In addition to your answers, UNAB will provide the Social Bee Lab with administrative records of your courses and your university entrance exam score. Your records, decisions, and your identity will be kept strictly confidential. Data about you collected within the scope of the study are used for scientific purposes only and are treated as strictly confidential. The Social Bee Lab will anonymize your data, and the researcher will analyze it without knowing your identity. All data generated will be stored on the researcher's computer. You have the right to access your personal data and request its deletion. You can exercise this right by contacting the researcher.

If something is unclear or you have any questions, you can contact [omitted for anonymous review].

If you have questions about your rights as a participant, you can contact [omitted for anonymous review].

By continuing to the next screen, you agree to participate in this study.

Before you start, please answer these four questions.

What is your gender?

[Male, Female]

What is the socio-economic stratum to which your family belongs?

[Stratum 1 to Stratum 6]

What is your father's highest acquired level of education?

[Primary school, High school, Technical school, Undergraduate, Graduate, Postgraduate, Not applicable.]

What is your mother's highest acquired level of education?

[Primary school, High school, Technical school, Undergraduate, Graduate, Postgraduate, Not applicable.]

Part 1

You will now participate in two quizzes, each lasting five minutes. Please try to answer them to the best of your ability.

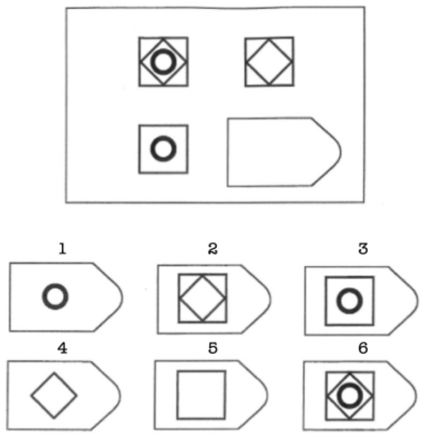
We will allocate up to [classroom-specific number equal to 20% of class size] bonuses of 100.000 pesos in this first part. The steps to allocate the bonuses for Part 1 are explained below.

[classroom-specific illustrations explaining the incentive structure]

[random assignment to either cognitive or social skills test]

Test - Cognitive Skill

In this test, you will see a series of images. Below is an example of the images you will solve. At the top of each image, there is a pattern with a piece that has been removed. Your task is to choose which of the six pieces completes the pattern correctly. For each image, there is only one correct piece. Look at the following example:



First, notice a square in the upper left, the upper right, and the lower left. Also, notice that the circle is eliminated when one moves from the upper left to the upper right. Finally, the rhombus is eliminated when moving from the upper left to the lower left. Therefore, the correct piece should eliminate the circle and the rhombus, leaving only a square. So, the correct answer is piece 5.

To give your answer to each image, you must choose the correct option and then continue to the next screen. After giving your answer you cannot go back.

You will have 5 minutes to complete the test, which consists of 18 images to solve. The percentage of correct answers will determine your chances of winning one of the 100.000 pesos bonuses if you are chosen for the drawing.

Are you ready?

Your 5 minutes will start as soon as you move to the next screen.

Problem 1

[screenshot of Raven's matrix]

[After participants submit an answer, a new matrix appears on the screen. The sequence of matrices is the same for all participants. Participants cannot return to a previous screen. Participants do not have to provide answers for all 18 matrices.]

You have finished the test. You can proceed to the next screen.

How did you do on the test?

If we randomly choose 10 participants from this classroom, how many people do you think solved fewer correct problems than you?

[Slider from 0 to 10]

Test - Emotions

In this test, you will see a series of photographs. Below is an example of the pictures you will see. In each picture, you will see the eyes of a person. Below the picture, you will see four possible emotions that this person is feeling. Your task is to choose which of the four emotions correctly describes what the person is feeling. For each picture, there is only one emotion. Look at the following example:

[Happy, Disappointed, Shocked, Worried]

In this case, the correct answer is: Shocked.

To give your answer to each picture, you must choose the correct option and then continue to the next screen. After giving your answer you will not be able to go back.

You will have 5 minutes to complete the test, which consists of 36 photographs to solve. The percentage of correct answers you get will determine your chances of winning one of the 100.000 pesos bonuses if you are chosen for the drawing.

Are you ready?

Your 5 minutes will start as soon as you move to the next screen.

Photograph 1: Choose the word that best describes the photograph

[photo from Multiracial Reading the Mind in the Eyes Test]

[After participants submit an answer, a new photo appears on the screen. The sequence of photos is the same for all participants. Participants cannot return to a previous screen. Participants do not have to provide answers for all 36 photos.]

You have finished the test. You can proceed to the next screen.

How did you do on the test?

If we randomly choose 10 participants from this classroom, how many people do you think solved fewer correct photographs than you?

[Slider from 0 to 10]

Part 2

At the beginning of this study, all participants took two tests, one on cognitive ability and one on emotions. In this part, we will ask you to recommend the people who in your opinion will score the best on each test.

You may recommend 3 people per test, but you may not recommend yourself.

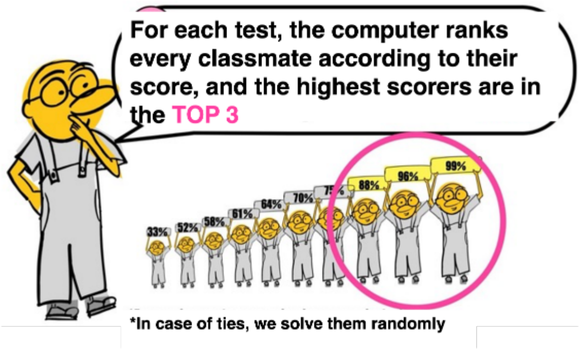
We will allocate up to [classroom-specific number equal to 40% of class size] bonuses of 100.000 pesos for Part 2. The steps for allocating bonuses are explained below.

[random assignment to either quota or baseline condition]

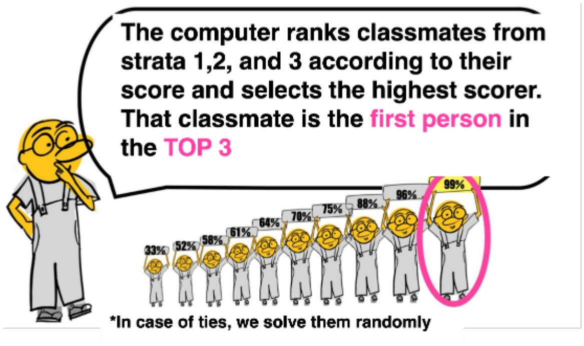
[classroom-specific illustrations explaining the incentive structure depending on assignment to either baseline or quota conditions]

[random assignment to either cognitive or social skills referral task]

Figure B.1: Illustrations for the two conditions



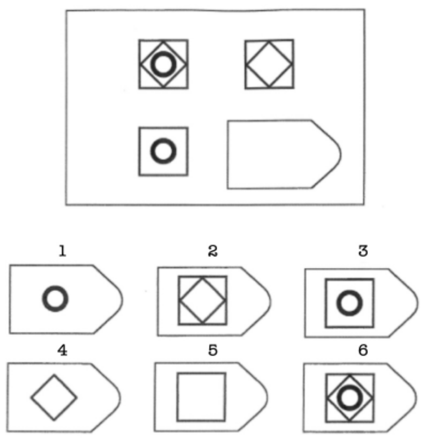
(a) Baseline



(b) Quota

Recommendation - Cognitive Skill

All participants took a test to identify the missing pattern in each image, as in the example below. This test is used to measure general intelligence.



Next, we will present you with a list of the names of all the students in this room. We will ask you to recommend the three people you think will score the highest on the general intelligence test.

If you are chosen by the computer, each of your recommendations in the top 3 increases your chances of winning one of the 100.000 pesos bonuses.

Select the students in this classroom who you consider to have the highest scores on the general intelligence test. (Select 3 students)

[Classroom-specific list of all classmate names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

Recommendation - Emotions

All participants took a test where they had to identify the emotion that best described the expression of each image as in the example below. This test is used to measure social skills.

Next, we will present you with a list of the names of all the students in this room. We will ask you to recommend 3 people you think will score the highest on the social skills test.

If you are chosen by the computer, each of your recommendations in the top 3 increases your chances of winning one of the 100.000 pesos bonuses.

Select the students in this classroom who you consider to have the highest scores on the social skills test. (Select 3 students)

[Classroom-specific list of all the names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

Part 3: Recommendation - Random draw

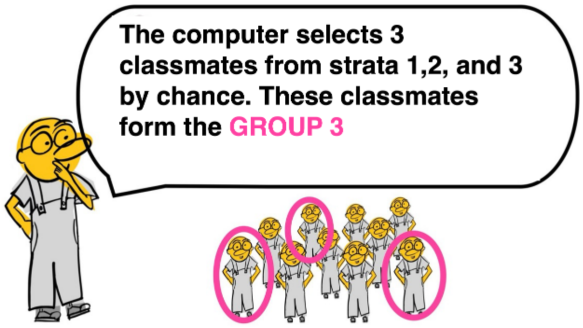
In this part, the computer will randomly choose three students who belong to strata 1, 2, or 3. We will ask you to nominate three people you think the computer will choose.

We will allocate up to [classroom-specific number equal to 20% of class size] bonuses of 100.000 pesos for Part 3. The steps for allocating the bonuses are explained below.

[classroom-specific illustrations explaining the incentive structure]

Select the students in this classroom who belong to strata 1, 2, or 3, who you think will be randomly selected by the computer (Select 3 students).

Figure B.2: Illustration for the Guessing Task



[Classroom-specific list of all the names visible on one screen. Participants have to pick 3 classmates to continue. Picking their own name invalidates their choices.]

Part 4

Do you want to know your scores on the general intelligence test and the social skills test? We can analyze the data and give you a report that explains your strengths in these two areas. Also, what do these strengths mean, and how can you leverage them for your personal and professional development?

If you want to receive your skills report, we need to contact you again. We also want to be able to invite you to new studies where you can participate for more bonus money. Please indicate if you agree to be contacted again.

[I can be contacted for new studies and to send me my report. I can be contacted to send my report, but not for new studies. No, I do not want to be contacted again.]

[if participant gives consent to be contacted again]

Please enter your contact email:

[student email]

Chapter 4

When Proximity Isn't Enough: Network Segregation and Class Bias in Referrals

Status: Not sent out for publication

Contribution Statement: Manu Munoz: Resources, Investigation, Software, Conceptualization, Methodology, Supervision, Funding acquisition, Writing - Review & Editing, Project administration. Ernesto Reuben: Conceptualization, Methodology, Supervision, Funding acquisition, Project administration. Reha Tuncer: Conceptualization, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization.

Abstract

The share of high-socioeconomic status (SES) connections in one's network is a strong correlate of labor market income. While universities provide ample opportunities for cross-SES contact, it remains unclear whether this exposure translates into meaningful connections. We investigate this question by exploring SES biases in referral selection. We conduct a lab-in-the-field experiment with 734 Colombian university students who make incentivized referrals from their enrollment networks. Randomizing participants between performance-only incentives and performance plus a fixed bonus for referral recipients, we find that referrals go to high-performing peers with whom they take many courses together, regardless of condition. While low-SES referrers exhibit strong in-group preferences, middle- and high-SES referrers show no biases toward their own and other groups. Network segregation, driven by cost-based program selection, limits cross-SES referral opportunities even without an explicit SES bias. Our results imply that institutional policies promoting cross-SES contact are key for reducing SES-based inequalities.

JEL Classification: C93, J71, D85, Z13

Keywords: inequality, economic mobility, peer networks, class discrimination, homophily

4.1 Introduction

Equally qualified individuals face different labor market outcomes based on their SES (Stansbury & Rodriguez, 2024). This persistent inequality undermines meritocratic ideals and represents a substantial barrier to economic mobility. A key driver of SES-based inequality in the labor market stems from differences in social capital.[1] Economic connectedness, defined as the share of high-SES connections in one's network, is an important facet of social capital because it correlates strongly with labor market income (Chetty et al., 2022a). In this sense, a lack of social capital means lack of access to individuals with influential (higher paid) jobs and related job opportunities. It implies having worse outcomes when using one's network to find jobs conditional on the capacity to leverage one's social network.[2]

Research on economic connectedness has focused on two distinct mechanisms that shape cross-SES connections: network composition (who you have the chance to meet inside an institutional environment) versus individual preference (who you choose to connect with among those available). A prevailing hypothesis is that increasing exposure to high-SES individuals will lead to higher rates of cross-SES connections in networks (Chetty et al., 2022b). Universities, in this regard, represent a particularly promising setting as they attract higher-than-population shares of high-SES students, and create more opportunities for cross-SES connections. However, whether these cross-SES connection opportunities turn into meaningful contacts, and the role of SES biases in the process has not yet been explored.

We address these questions through a referral experiment in a university setting. Focusing on the role of SES in referral selection, we studied whether individuals tended to refer same-SES peers. We recruited 734 undergraduate students to make incentivized referrals among peers they encountered during their coursework. Referrals were made for the math and critical reading areas of the national university entry exam. To incentivize performance-based referral selection, participants (referrers) earned payments up to $60 per referral based on their nominee's percentile ranking at the university. This setup provided an objective performance benchmark for referrals where SES biases in referral selection could still play a role.

[1]See for example Bourdieu (1986); Loury (1977) for pioneering work on the relationship between social position and human capital acquisition.

[2]See for example Lin et al. (1981); Mouw (2003) for differential outcomes while using contacts in job search, and Pedulla and Pager (2019); Smith (2005) specifically for the effects of race conditional on network use.

Referrals originated from each participant's unique course enrollment network that we constructed using extensive administrative data. The enrollment network covered each course that the participant had taken with all other undergraduate students at the university (more than 4,500 individuals). It allowed us to observe every potential referral candidate, and the tie strength between the candidate and the referrer, which we measured by the number of courses they took together. Referrals from the enrollment networks enabled us to separate network composition (i.e., chance of meeting during coursework and frequency of contact) from SES biases in referral selection (i.e., individual choice in picking a referral). By doing so, we were able to control for naturally varying network compositions with referral candidates at the individual level, and could identify group-level SES biases in referral selection that go beyond mere opportunities to interact at the university.

We randomized participants into two conditions. In the **Baseline** condition, participants made referrals with performance-based incentives only, where their earnings depended on the actual performance of their referrals. In the **Bonus** condition, participants made referrals with performance-based incentives and an additional fixed bonus ($25) going to their referral of choice. The fixed bonus created incentives to refer peers even if they performed less well, potentially amplifying the relevance of other factors like the SES bias and tie strength.

We find that referrals consistently go to higher-performing peers with high tie strength, regardless of conditions and exam areas. Pooling across these, we find that SES bias in referral selection is primarily driven by low-SES participants exhibiting in-group preferences. In our preferred specification, low-SES referrers are 27% more likely to refer other low-SES peers and 36% less likely to refer high-SES relative to middle-SES peers. In contrast, middle- and high-SES referrers show no biases toward their own or other groups.

With 93% of referrals going to peers within the same academic program with whom referrers have taken many courses together, we find that network composition rather than SES biases better explains the observed referral patterns. At the tie strength where referrals typically occur (median 12 courses together), network segregation becomes stark: low-SES students comprise 44.5% of low-SES referrers' networks compared to only 15.7%

of high-SES referrers' networks, despite representing 34% of the university population. This segregation means that even without bias against low-SES peers, high-SES referrers rarely encounter low-SES candidates among their close university connections.

Looking for potential mechanisms driving the segregation in enrollment networks, we identify program selection as key. Program fees at our partner university are fixed on a cost basis, and less than 5% of undergraduates qualify for scholarships. One consequence of these policies is that SES groups end up sorting into programs on the basis of their costs, where some programs cost up to six times more on a yearly basis. To sum, even though low-SES are exposed to higher-than-population shares of high-SES students, and high-SES are not biased toward other SES groups, meaningful interaction opportunities at the university are genuinely limited.

Our findings should be interpreted with some scope conditions. First, our referrals have no direct job consequences, and participants refer under anonymity. These may represent a lower stake environment for referrers. Nevertheless, we replicate typical findings from earlier referral experiments where performance-based incentives brings in qualified candidates from referrer networks (e.g., Beaman and Magruder (2012); Witte (2021)).

Second, enrollment networks capture classroom-based interactions and their intensity rather than broader networks of close friendships. Unlike self-reported friendship networks that must limit the number of connections surveyed (Griffith, 2022), administrative data captures all classroom interactions without artificial size constraints. Combining enrollment networks with additional network data (e.g., from social media) could be useful for better identifying interactions at the university. Still, we find that tie strength predicts referral selection well beyond same program affiliation, suggesting it does capture meaningful variation in social interactions.

Finally, our setting examines SES bias within a single institution where cross-SES contact is possible, and the networks of different SES groups are separated due to program selection. The generalizability to contexts with different institutional structures remains an open question for future research.

We contribute to several strands of literature. First, a burgeoning literature studies the effects of SES on labor market outcomes (Friedman & Laurison, 2019; Laurison & Friedman, 2024; Stansbury & Rodriguez, 2024), with mechanisms including cultural matching and SES-based discrimination in the hiring processes (Galos, 2024; Núñez & Gutiérrez, 2004; Rivera, 2012; Rivera & Tilcsik, 2016). We extend this literature by examining the role of referral networks as a specific mechanism through which SES could affect economic opportunities.

A subset of the literature focuses on SES-based differences in social capital and network formation (Chetty et al., 2022a; Engzell & Wilmers, 2025; Michelman et al., 2022), with tie strength (Gee et al., 2017; Kramarz & Skans, 2014; Sterling, 2014; Wang, 2013) and homophily (Bolte et al., 2024; Currarini et al., 2009; Jackson, 2022; McPherson et al., 2001; Montgomery, 1991) driving differences across groups. Based on the pioneering work of Currarini et al. (2010), we contribute by identifying two different types of homophily, and separate whether differential referral outcomes stem from network composition (who you know) versus taste-based biases (who you choose to interact with). Our findings suggest that structural factors impacting network composition, rather than taste-based SES biases, drive the differences in referral outcomes. Under this light, implementing mixed-program courses to increase across-SES tie strength should be a clear policy goal in order to reduce SES-based network segregation.

Third, we contribute to the literature on job referral experiments. This literature provides causal evidence on why referrals in the labor market are prevalent,[3] finding that performance-based incentives bring in qualified candidates otherwise not identified by demographic characteristics (Beaman & Magruder, 2012; Friebel et al., 2023; Pallais & Sands, 2016; Witte, 2021), and the consequences of relying upon referral hiring, which come at the cost of disadvantaging certain groups (Beaman et al., 2018; Hederos et al., 2025). We extend this literature by causally evaluating the effects of a sizeable monetary bonus for referral candidates and exploring SES biases in referral selection.

The remainder of the paper is organized as follows. Section 4.2 begins with the background and setting in Colombia. Section 4.3 presents the empirical strategy and Section

[3] Referrals solve frictions in the search and matching process and benefit both job-seekers and employers (Topa, 2019). Referral candidates tend to get hired more often, have lower turnover, and earn higher wages (Brown et al., 2016; Dustmann et al., 2016; Obukhova & Lan, 2013).

4.4 presents the design of the experiment. In Section 4.5 we describe the experimental sample, incentives and the procedure. Section 4.6 discusses the results of the experiment and Section 4.7 discusses potential mechanisms and robustness checks. Section 4.8 concludes. The Appendix presents additional tables and figures as well as the experiment instructions.

4.2 Background and Setting

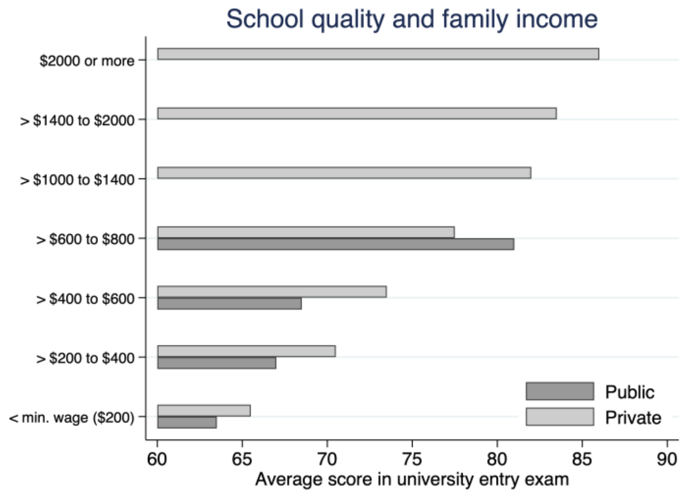
4.2.1 Inequality and SES in Colombia

Our experiment took place in Colombia, a country that consistently ranks highly in terms of economic inequality. The richest decile of Colombians earn 50 times more than the poorest decile (United Nations, 2023; World Bank, 2024). This economic disparity creates profound differences in outcomes across SES groups in terms of education, geographic residence, language, manners, and social networks (Angulo et al., 2012; García et al., 2015; García Villegas & Cobo, 2021).

In higher education, economic inequality manifests itself by preventing meaningful interaction between SES groups. Wealthy families attend exclusive private schools while poorer families access lower-quality public or "non-elite" private institutions (see Figure 4.1). While similar patterns also exist elsewhere, differences in educational outcomes across SES groups are particularly visible in Colombia.

We rely on Colombia's established estrato classification system to measure SES in our study. In 1994, Colombia introduced a nationwide system that divides the population into six strata based on "similar social and economic characteristics" (Hudson & Library of Congress, 2010, p. 102). Designed for utility subsidies from higher strata to support lower strata, the system aligns with and reinforces existing social class divisions (Guevara S & Shields, 2019; Uribe-Mallarino, 2008). It is also widely used by policymakers and in official statistics (Fergusson & Flórez, 2021a). Using the estrato system, we categorize students in strata 1-2 as low-SES, strata 3-4 as middle-SES, and strata 5-6 as high-SES.

Figure 4.1: Income, performance, and university choice in Colombia



Note: This figure shows the average score national university entry exam by monthly family income and type of higher education institution. With average student scores in the 65-70 band, the private university where we conducted this study caters to both low- and high-income students. Figure reproduced from Fergusson and Flórez (2021b).

4.2.2 Partner institution and the enrollment network

Our study takes place in a non-elite private university which attracts students across the socioeconomic spectrum: the university's undergraduate student body comprises 34% low-SES, 51% middle-SES, and 15% high-SES students.[4] This diversity provides opportunities for different SES groups to meet and interact within the same institutional framework.

The contact at the university is on equal status. All undergraduate students pay the same fees based on their program choices, and less than 5% of undergraduate students receive scholarships. The student body is mostly urban (> 70%), not part of an ethnic minority (> 95%), and has similar university entry exam scores (see Appendix Figures A.1a and A.1b). These make our setting appropriate to study the effects of contact on intergroup discrimination.

[4]Government statistics reveal less than 5% of the population is high-SES (Hudson & Library of Congress, 2010, p. 103).

Undergraduate students at the university choose among 32 different academic programs. Students take between 5 and 7 courses per semester, and programs last between 4 and 12 semesters (2 to 6 years). The majority (64%) of students are enrolled in the 10 programs described in Appendix Figure A.2. While medicine, the largest program by size at the university lasts for 12 semesters, specialized programs for immediate entry into the workforce last only 4 semesters. Academic program choice thus shapes students' connections at the university, influencing both who they encounter in classes and the frequency of these interactions.

To map these social connections, we construct enrollment networks using administrative data. For each participant, we identify all other undergraduate students with whom they have taken at least one course and create their individual network of university connections. The size of this network depends on how many students a participant has encountered through coursework, while the tie strength is measured by the number of courses taken together. This approach provides a complete picture of each participant's social environment at the university, and includes detailed characteristics (i.e., SES, academic program, performance) for both the participant and every person in their network.

4.3 Empirical Strategy

We use a conditional logit model to study SES biases in referral selection. Our dependent variable follows a multinomial distribution where referrer i selects one candidate j from their enrollment network for two exam areas. For each referrer, we observe all potential candidates, i.e, students they took at least one course with, along with their characteristics. The conditional logit model with individual fixed effects takes the form:

$$Y_{ij} = \alpha_i + \beta_1 SES_{ij} + \beta X_{ij} + \varepsilon_{ij} \tag{4.1}$$

where $Y_{ij} = 1$ if referrer i chooses referral candidate j, and 0 otherwise. We set middle-SES as the base category, so β_1 is the log-odds estimate for referring low- and high-SES candidates relative to middle-SES. X_{ij} includes the remaining characteristics of referral candidates in the enrollment network that improve model fit such as entry exam scores and the number of courses taken together with the referrer. These continuous variables are standardized using means and standard deviations calculated by first computing network-level statistics for each referrer, then averaging across all 734 networks.[5] The individual fixed effects α_i control for referrer-specific factors that might influence both network formation and referral decisions. Because we observe two referrals (one per exam area) from each referrer, we cluster standard errors at the referrer level and account for the potential correlation in the error terms.

The key advantage of this approach is that by conditioning on each referrer's enrollment network, we eliminate selection bias from program choice and other factors that determine who appears in each person's choice set. The identifying variation comes from within-network differences in referral decisions, holding constant the pool of available candidates. We estimate separate models for each referrer SES group to estimate aggregate SES biases across socioeconomic groups.

For identification, we require two assumptions. First, conditional exogeneity. SES and the number of courses taken together could be endogenous due to program selection. High-SES students may sort into expensive programs while low-SES students choose affordable programs, creating SES variation across enrollment networks. Similarly, the number of courses taken together reflects program selection decisions that may correlate with unobserved referral preferences. However, conditional on the realized enrollment network, the remaining variation in both SES and the number of courses taken together across referral candidates must be independent of unobserved factors affecting referral decisions. As a robustness check, we show that being in the same program with the referrer does not impact our SES bias estimates, although it reduces the coefficient estimate for the number of courses taken together.

Second, the independence of irrelevant alternatives. This assumption could be violated if peers within the same SES group are viewed as close substitutes, where adding similar

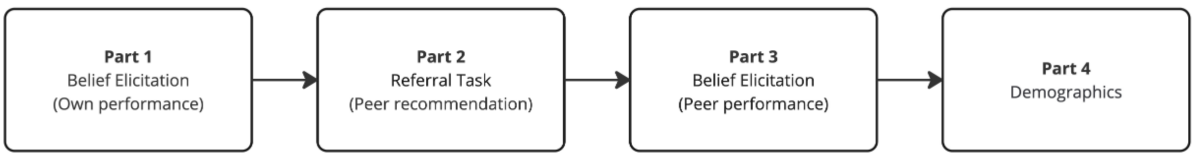
[5]Each referral candidate's entry exam score and the number of courses they have taken with the referrer is standardized using these sample-level statistics. The standardization formula is $z_i = (x_i - \bar{X})/\sigma$, where $\bar{X}$ and σ are the average mean and standard deviation across participant networks for the measure.

alternatives distorts choice probabilities. While this concern may have some validity in our setting,[6] alternative discrete choice models that relax IIA are computationally prohibitive given our large dataset.[7] We therefore proceed with the conditional logit framework while acknowledging its limitations.

4.4 Design

We designed an experiment to assess SES biases in referral selection and evaluate the causal effect of providing bonuses to referral candidates. The 30-minute experiment consisted of three sequential tasks: initial belief elicitation about participants' own performance on the national university entry exam, referral tasks where they nominated peers for two exam areas (math and critical reading), and another belief elicitation about their nominees' performance. This structure allowed us to collected incentivized measures for the accuracy of participants' performance beliefs and their referral decisions. Figure 4.2 shows the experiment overview, and detailed instructions are provided in Appendix A.2.

Figure 4.2: Experiment Overview



Note: Participants first reported beliefs about their university entry exam performance in two areas, then made referrals for those, and finally reported beliefs about their referrals' performance and provided demographics.

[6] Among participants making referrals to two different individuals, half refer to someone else from the same SES, suggesting potential substitutability within SES groups.

[7] Models such as nested logit become computationally intractable with over 250,000 observations across 734 individuals.

4.4.1 Performance measures

To establish an objective basis for referral performance, we use national university entry exam scores (SABER 11). All Colombian high school students take the SABER 11 exam at the end of their final year as a requirement for university admission. The scores from this exam provide pre-existing, comparable measures of performance.

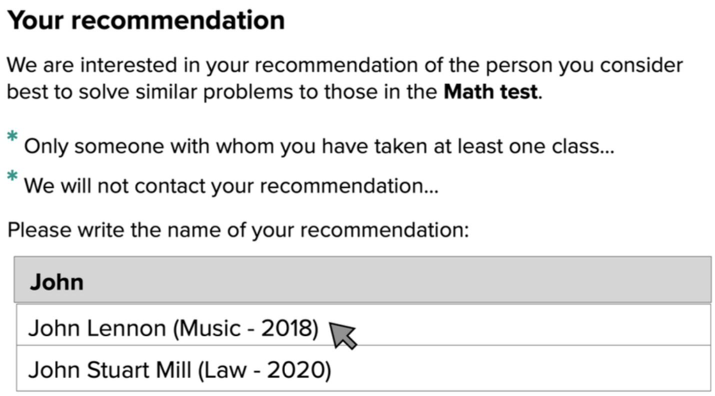
The exam consists of five areas (critical reading, mathematics, natural sciences, social sciences, and English). We focus on critical reading and mathematics as these represent independent and overarching skills. Critical reading evaluates competencies necessary to understand, interpret, and evaluate texts found in everyday life and broad academic fields (e.g., history). Mathematics assesses students' competency in using high school level mathematical tools (e.g., reasoning in proportions, financial literacy). These together capture performance in comprehending and critically evaluating written material as well as reasoning and problem-solving abilities.

For each area, we calculate percentile rankings based on the distribution of scores among all currently enrolled students, providing a standardized measure of relative performance within the university population.

4.4.2 Referral task

The main task involves making referrals among peers. For both exam areas (critical reading and mathematics), participants refer one peer they believe excels in that area. We provide an example question from the relevant exam area to clarify the skills that are being assessed. Participants type the name of their preferred candidate to make a referral. To avoid issues with recall, the interface provides autocomplete name and program suggestions from the administrative database (see Figure 4.3).

Figure 4.3: Referral task interface



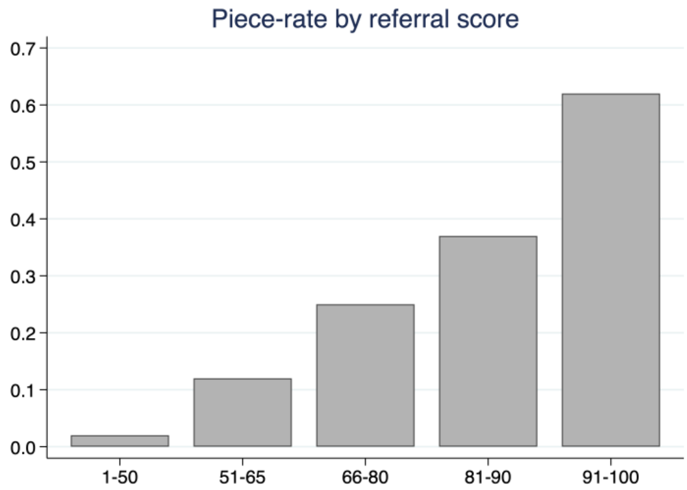
Note: This illustration shows how the system provides suggestions from enrolled students with their program and year of study from the administrative database.

Participants are required to only refer students with whom they have taken at least one class during their university studies. Referrals to students without any classes taken together are possible, but not valid. This condition ensures that referrals are based on actual peer interactions. We randomize the order in which participants make referrals across the two exam areas.

We incentivize referrals using a piece rate payment structure. Referrers earn increasing payments as the percentile ranking of their referral increases (see Figure 4.4). We multiply the piece rate coefficient associated with the percentile rank by the actual exam scores of the referral to calculate earnings. This payment structure provides strong incentives to refer highly ranked peers with potential earnings going up to $60 per referral. [8]

[8]Note that due to the selection into the university, the actual exam score distribution has limited variance. Below a certain threshold students cannot qualify for the institution and choose a lower ranked university, and above a certain threshold they have better options to choose from.

Figure 4.4: Referral incentives



Note: This figure shows how the piece rate coefficient increases as a function of the referral ranking in the university, providing incrementally higher rewards for higher ranked peers.

4.4.3 Bonus Treatment

To examine how different incentive structures affect referral selection, we randomly assign a fixed bonus payment for students who get a referral (receiver). In the **Baseline** treatment, only the participants, i.e., those who make referrals (sender), can earn money based on their referral's performance. The **Bonus** treatment adds a fixed payment of $25 uniquely to the peer who gets the referral. This payment is independent of the referral's actual performance (see Table 4.1).

Table 4.1: Incentive structure by treatment

	Baseline	**Bonus**
Referrer (sender)	Performance-based	Performance-based
Referral (receiver)	No payment	Fixed reward

We use a between-subjects design and randomly assign half our participants to the **Bonus** treatment. This allows us to causally identify the effect of the bonus on referral selection. Participants learn whether their referral gets the fixed bonus before making referral decisions.

4.4.4 Belief elicitation

We collect two sets of beliefs to assess the accuracy of participants' knowledge about exam performance. Participants first report beliefs about their own percentile ranking in the university for math and critical reading areas. After making referrals, participants report their beliefs about their referrals' percentile ranking in the university. For both belief elicitation tasks, participants earn \$5 per correct belief if their guess is within 7 percentiles of the true value. This margin of error is designed to balance precision with the difficulty of the task.

4.5 Sample, Incentives, and Procedure

We invited all 4,417 undergraduate students who had completed their first semester at the university at the time of recruitment to participate in our experiment. A total of 837 students participated in the data collection (19% response rate). Our final sample consists of 734 individuals who referred peers with whom they had taken at least one class together, excluding 12% of participants who made two non-valid referrals.

Table 4.2 presents key demographic characteristics and academic performance indicators across treatments (see Appendix Table A.1 for selection). The sample is well-balanced between the **Baseline** and **Bonus** conditions, and we observe no statistically significant differences in any of the reported variables (all p values > 0.1). Our sample is characterized by a majority of middle-SES students with about one-tenth of the sample being high-SES students. The test scores and grade distributions are balanced. On average, participants had taken 3.8 courses together with members of their network, and the average network consisted of 175 peers.

Table 4.2: Balance between treatments

	Baseline	**Bonus**	***p***
Reading score	64.712	65.693	0.134
Math score	67.366	67.597	0.780
GPA	4.003	4.021	0.445
Connections	173.40	176.88	0.574
Courses taken	3.939	3.719	0.443
Low-SES	0.419	0.401	0.615
Middle-SES	0.492	0.506	0.714
High-SES	0.089	0.094	0.824
Observations	382	352	734

Note: This table presents balance tests between **Baseline** and **Bonus** conditions. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests with unequal variances. All reported p-values are two-tailed. Reading and math scores are in original scale units out of 100. GPA is grade point average out of 5. Connections refer to the average number of network members. Low-SES, Med-SES, and High-SES indicate SES categories based on strata.

The experiment was conducted online through Qualtrics, and we recruited participants by sending invitations to their student emails. To ensure data quality while managing costs, we randomly selected one in ten participants for payment. Selected participants received a fixed payment of \$17 for completion. They also received potential earnings from one randomly selected belief question (up to \$5) and one randomly selected referral question (up to \$60). This structure resulted in maximum total earnings of \$82. The average time to complete the survey was 30 minutes, with an average compensation of \$80 for one in ten participants randomly selected for payment. Payment processing occurred through bank transfer within 15 business days of participation.

4.6 Results

4.6.1 Network characteristics

We begin by describing the key features of the enrollment networks. On average, participants connect with 175 other students, and take an average of 3.8 courses together. Figure 4.5 shows how network characteristics vary by students' time at the university: both the number of connections (network size) and the number of courses taken together (tie strength) change as participants progress through their studies.

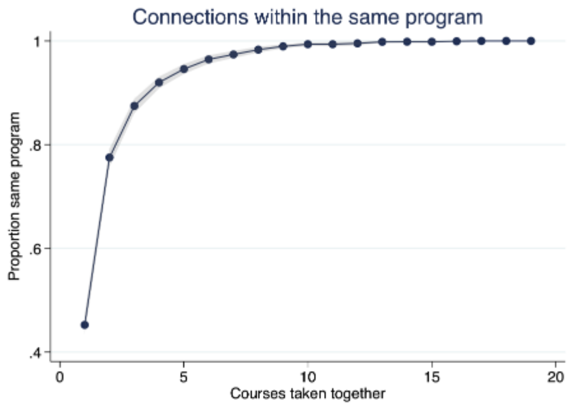
Figure 4.5: Network size and courses taken together by time spent at the university

Note: This figure displays the average number of connections in gray and the average number of courses taken together with connections in blue across semesters completed. Network size (nb. of connections) peaks around 7 semesters before declining as students graduate. Tie strength (nb. of courses taken) has an increasing trend.

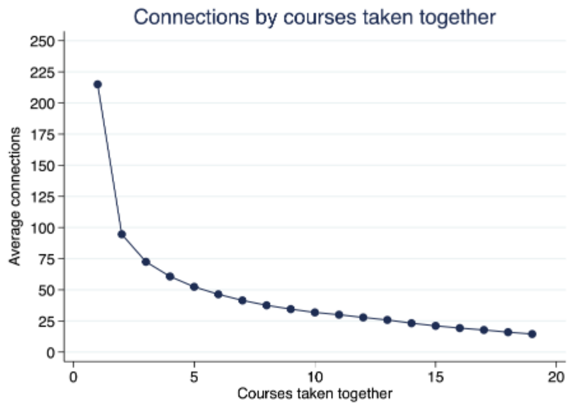
We now examine how tie strength relates to network size and composition. First, if two students take more courses together, it is very likely that they are in the same academic program. We plot this relationship in Figure 4.6a: As students take more than 5 courses together, the share of students in their enrollment network from the same academic program quickly exceeds 90%. Second, because students sort into specialized academic programs, increases in courses taken together should result in decreases in connections.

We plot this relationship in Figure 4.6b: As students take more than 5 courses together, the size of their enrollment network drops dramatically from above 210 to below 50. These patterns reveal that while participants' overall networks are large with relatively few courses taken together on average, they are more frequently in contact within a much smaller group of peers from the same academic program.

Figure 4.6: Network characteristics and courses taken together



(a) Same-program share



(b) Network size

Note: Panel (a) illustrates the share of connections within the same program as a function of the number of courses taken together. Panel (b) shows the average network size as a function of the number of courses taken together. Taking more than 5 courses together with a network member means on average 90% chance to be in the same program. Similarly, past 5 courses together, the average network size dwindles by 80%, from more than 210 individuals to below 50.

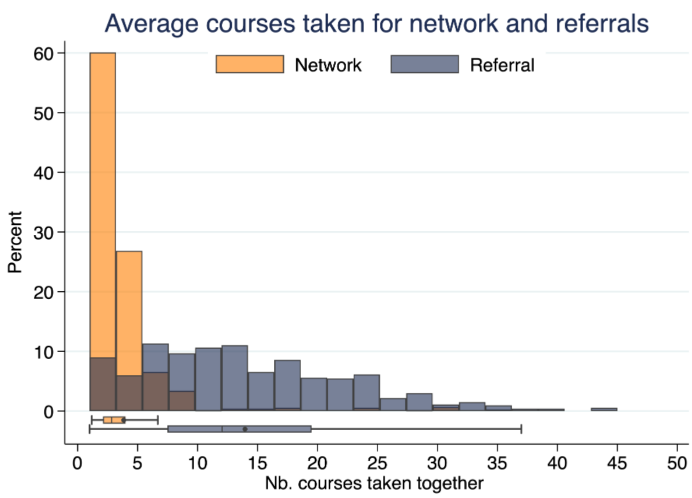
4.6.2 Referral characteristics

Participants made one referral for math and one referral for the reading part of the university entry exam from their enrollment networks. We collected 1,342 valid referrals from 734 participants in our final dataset. More than 90% of these consist of participants referring for both exam areas (see Appendix Table A.2). About 70% of these referrals go to two separate individuals. We compare the outcomes across exam areas for referrals only going to separate individuals in Appendix Table A.3 and all referrals in Appendix Table A.4. In both cases, we find no meaningful differences between referrals made for math or critical reading areas of the entry exam. As referrals in both exam areas come from the same enrollment network, we group referrals per participant and report average outcomes.

What are the characteristics of the individuals who receive referrals, and how do they compare to others in the enrollment network? Because we have an entire pool of potential candidates with one referral chosen from it, we compare the distributions for our variables of interest between the referred and non-referred students.

First, referrals go to peers with whom the referrer has taken around 14 courses with on average, compared to almost 4 on average with others in their network (see Figure 4.7). This difference of 10 courses is significant ($t = 34.98$, $p < 0.001$), indicating that referrers choose individuals with whom they have higher tie strength. While the median referral recipient has taken 12 courses together with the referrer, the median network member has shared only 2.8 courses. The interquartile range for referrals spans from 7.5 to 19.5 courses, compared to just 2.1 to 4.0 courses for the broader network, highlighting the concentration of referrals among peers with higher tie strength. In addition, 93% of referrals go to students in the same program as the referrer.

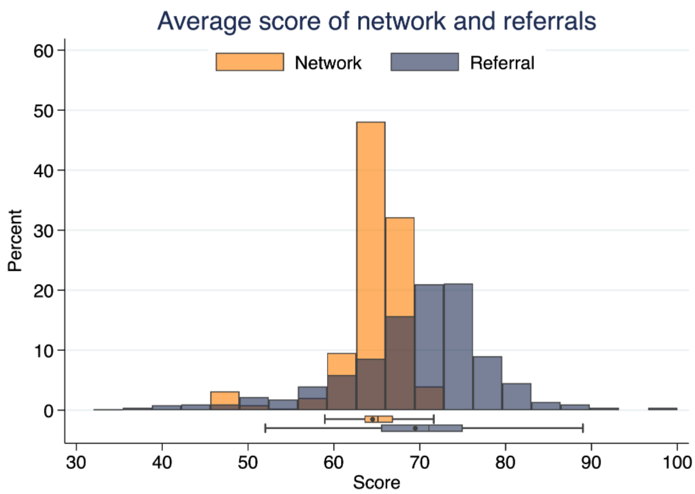
Figure 4.7: Courses taken together with network members and referrals



Note: This figure compares the distributions of the number of courses taken together between referrers and their network members (orange) versus referrers and their chosen referral recipients (dark blue) for all 734 participants. 75% of referral recipients take more than 7.5 courses together with the referrer, compared to only 25% of network members. The distributions are significantly different (Kolmogorov-Smirnov test $D = 33.37$, $p < 0.001$).

Second, we examine entry exam score differences between referred students and the broader network. Referrals go to peers with an average score of 69.5 points, compared to 64.5 points for other network members (see Figure 4.8). This difference of 5 points is significant ($t = 18.97$, $p < 0.001$), indicating that referrers choose higher-performing peers. While the median referral recipient scores 71 points, the median network member scores 65.1 points. The interquartile range for referrals spans from 65.5 to 75 points, compared to 63.5 to 66.9 points for the broader network, highlighting the concentration of referrals among higher performing peers. Participant beliefs regarding their referral's and own performance rank at the university also support these findings (see Appendix Figures A.3a and A.3b).

Figure 4.8: Entry exam scores of network members and referrals



Note: This figure compares the distributions of entry exam scores (math and critical reading average) between referrers' network members (orange) versus their chosen referral recipients (dark blue) for all 734 participants. 75% of referral recipients score above 65.5 points compared to only 25% of network members scoring above 66.9 points. The distributions are significantly different (Kolmogorov-Smirnov test $D = 71.16$, $p < 0.001$).

4.6.3 Effect of the Bonus treatment

Do referrals across treatments have different outcomes? We compare the performance and the number of courses taken together with the referrer between the **Baseline** and **Bonus** treatments in Table 4.3. Contrary to our expectations, we find that the number of courses taken together with referrer, as well as performance measures across critical reading and math (including grades) are similar across treatments. Taken together, the results on academic performance and tie strength suggest these two factors drive referrals regardless of treatment. For this reason, in the remainder of the paper, we report pooled results combining the averages of referral outcomes across treatments.

Table 4.3: Characteristics of referrals by treatment

	Baseline Referred	**Bonus Referred**	***p***
Reading score	67.806	67.210	0.308
Math score	70.784	70.155	0.406
GPA	4.155	4.149	0.799
Courses taken	13.840	14.065	0.723
Observations	382	352	

Note: This table compares the characteristics of network members who were referred under baseline vs. bonus treatments. p-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample t-tests with unequal variances. All reported p-values are two-tailed. Reading and math scores are raw test scores out of 100. GPA is grade point average out of 5. Courses taken is the number of courses participant has taken with their referral. Both columns only include the average outcomes of network members who were referred in each treatment.

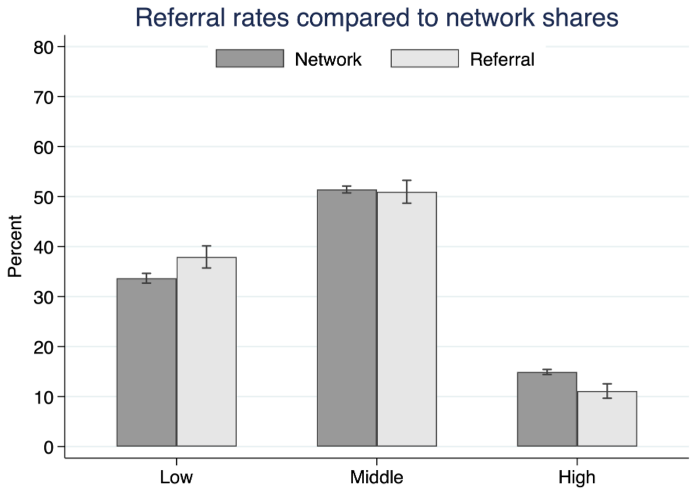
4.6.4 Referral SES composition

To motivate the SES biases in referral selection, we now examine the overall SES composition of referrals compared to the average network availability. Descriptively, referral patterns largely mirror underlying network structure.[9] Referrals to low-SES peers constitute 37.9% of all referrals compared to 33.7% network share, middle-SES referrals account

[9]Because we calculate the share of SES groups in every individual network, we get very precise estimates of the actual means. However, it is important to note that these are not independent observations. Each enrollment network is a draw with replacement from the same pool of university population, from which we calculate the proportion of SES groups per individual network, and take the average over an SES group. Pooling over SES groups who are connected with similar others systematically reduces variance (similar to resampling in bootstrapping). For this reason we choose not reporting test results in certain sections including this one and focus on describing the relationships between SES groups.

for 51.0% versus 51.4%, and high-SES referrals represent 11.1% compared to 14.9% (see Figure 4.9). The largest deviation is less than 5 percentage points for any SES group.

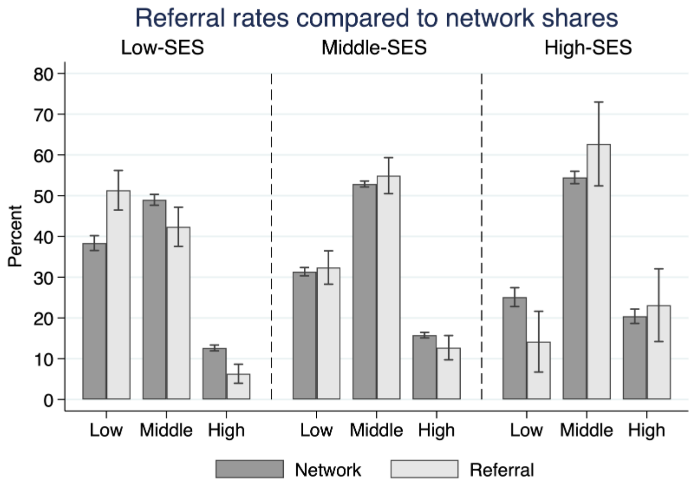
Figure 4.9: Referral patterns compared to network composition



Note: This figure compares the average SES composition of referrers' networks (dark gray) to the SES composition of referrals (light gray). Error bars represent 95% confidence intervals.

Examining patterns by referrer SES reveals larger deviations. Low-SES referrers have the largest same-SES deviation, referring 12.9 percentage points more to low-SES students than their network composition suggests, while high-SES referrers under-refer to low-SES students by 10.9 percentage points (see Figure 4.10). These descriptive findings suggest that referral selection in SES terms diverges most from underlying network structure when SES groups are further apart, and motivate our formal analysis.

Figure 4.10: Referral patterns by referrer SES compared to network composition



Note: This figure compares the average SES composition of referrers' networks (dark gray) to the SES composition of referrals (light gray) for low-, middle- and high-SES referrers (left to right). Error bars represent 95% confidence intervals.

4.6.5 Identifying the SES bias in referrals

We now describe our findings using the regression specification (see Equation 4.1) in Table 4.4. We first run three separate regressions, one for each referrer SES group, with a single regressor which is the referral candidate's SES. Controlling for network composition, we find that low-SES participants are more likely to refer other low-SES, and are less likely to refer high-SES relative to the probability of referring middle-SES peers. In contrast, we find that high-SES participants are less likely to refer other low-SES, relative to the probability of referring middle-SES peers.

Table 4.4: SES bias in referral decisions by referrer SES group

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.453***	-0.019	-0.710**
	(0.109)	(0.098)	(0.333)
High-SES referral	-0.584***	-0.255*	0.001
	(0.211)	(0.145)	(0.261)
χ^2	33.47	3.18	4.94
Observations	110,142	127,088	19,767
Individuals	301	366	67

Note: Individual-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a separate conditional logit regression estimated on the subsample of referrers from the indicated SES group. Coefficients represent log-odds of referring from the specified SES group relative to referring middle-SES candidates. All models include individual fixed effects that control for each referrer's choice set composition.

Next, we include a control for tie strength. We proceed by adding the standardized number of courses taken together as a control in our specification and describe the results in Table 4.5. A one standard deviation increase in the number of courses taken together proves to be highly significant across all models, with coefficients ranging from 0.856 to 1.049, indicating that tie strength substantially increases the probability of referral. The high χ^2 statistics suggest that the model with this regressor provides a better fit than previous models. We find that low-SES participants still show a strong same-SES bias relative to referring middle-SES peers at the average number of courses taken together. This same-SES bias is not observed among middle-SES or high-SES referrers, who also display no statistically significant bias toward low-SES candidates. No referrer group shows a positive bias for high-SES candidates relative to middle-SES candidates.

Table 4.5: SES bias in referral decisions by referrer SES group

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.348***	-0.064	-0.489
	(0.123)	(0.115)	(0.337)
High-SES referral	-0.366	-0.165	-0.140
	(0.223)	(0.157)	(0.286)
Courses taken (z-score)	0.856***	0.931***	1.049***
	(0.035)	(0.037)	(0.126)
χ^2	626.15	636.10	71.43
Observations	110,142	127,088	19,767
Individuals	301	366	67

Note: Individual-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a separate conditional logit regression estimated on the subsample of referrers from the indicated SES group. Coefficients represent log-odds of referring from the specified SES group relative to referring middle-SES candidates. All models include individual fixed effects that control for each referrer's choice set composition.

We then add standardized entry exam scores as a second control variable and describe our results in Table 4.6. A one standard deviation increase in the entry exam score (math and critical reading average) proves highly significant across all models, with coefficients ranging from 0.587 to 0.883. This shows merit-based considerations due to the incentive structure of the experiment remained central to referral decisions. The slightly higher χ^2 statistics compared to the earlier specification suggests that entry exam scores improve model fit. The inclusion of standardized entry exam scores strengthens SES biases: Low-SES referrers maintain their same-SES bias, with now a significant negative bias against high-SES. Middle-SES referrers, previously showing no SES bias, now show marginal negative bias against high-SES. Finally, high-SES referrers exhibit marginal negative bias against low-SES candidates.

The evidence of a bias becoming significant when controlling for entry exam scores has a nuanced interpretation. While at the university-level, low-SES typically score lower in the entry exam, low-SES students appearing in high-SES networks are positively selected, scoring about 0.14 standard deviations higher than middle-SES students (see Appendix Table A.5). Controlling for performance thus removes this positive selection and reveals the SES bias that was previously underestimated by above average performance of low-SES. Vice versa, high-SES in low-SES networks perform 0.12 standard deviations better than middle-SES students. The bias was underestimated as high-SES candidates' better performance relative to middle-SES increased referrals. Controlling for exam scores reveal that both high- and low-SES referrers have negative SES bias toward one another that operates independently of – and counter to – performance-based considerations. What makes a symmetric bias interpretation difficult is that while biased against low-SES, high-SES referrers do not (under any specification) display a positive bias toward their in-group.

We conclude that the SES bias in referral selection is primarily driven by low-SES referrers who exhibit strong in-group preferences. Middle- and high-SES referrers show no systematic discrimination against other SES groups once we account for network composition and other relevant factors contributing to the referral decision. We will next explore potential mechanisms that help explain the unexpected direction of the SES bias.

Table 4.6: SES bias in referral decisions by referrer SES group with academic performance controls

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.242**	-0.159	-0.600*
	(0.123)	(0.114)	(0.327)
High-SES referral	-0.445**	-0.274*	-0.345
	(0.222)	(0.157)	(0.287)
Courses taken (z-score)	0.859***	0.948***	1.043***
	(0.036)	(0.038)	(0.118)
Entry exam (referral z-score)	0.607***	0.587***	0.883***
	(0.052)	(0.047)	(0.111)
χ^2	789.87	756.06	120.54
Observations	110,142	127,088	19,767
Individuals	301	366	67

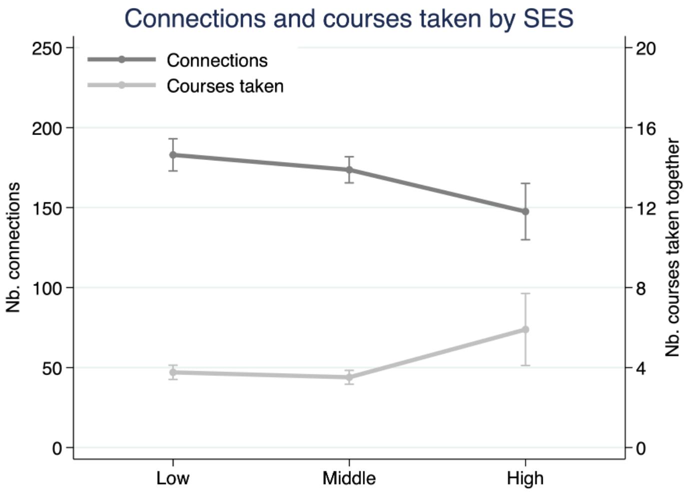
Note: Individual-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a separate conditional logit regression estimated on the subsample of referrers from the indicated SES group. Coefficients represent log-odds of referring candidates from the specified SES group relative to referring middle-SES candidates. All models include individual fixed effects that control for each referrer's choice set composition.

4.7 Potential Mechanisms and Robustness Checks

4.7.1 SES diversity in networks

How do enrollment networks differ across SES groups? We look at how the number of connections (network size) and number of courses taken together (tie strength) change across SES groups in Figure 4.11. Both low- and middle-SES students have significantly larger networks than high-SES students ($t = 3.03$, $p = 0.003$ and $t = 2.49$, $p = 0.013$, respectively), while high-SES students take significantly more courses with their network members than both low- ($t = -3.70, p < .001$) and middle-SES ($t = -4.20$, $p < .001$).

Figure 4.11: Network size and courses taken together by SES



Note: This figure displays the average number of connections and the average number of classes taken together across SES groups. The data shows a decrease in the number of connections with SES, and an associated increase in the number of classes taken together.

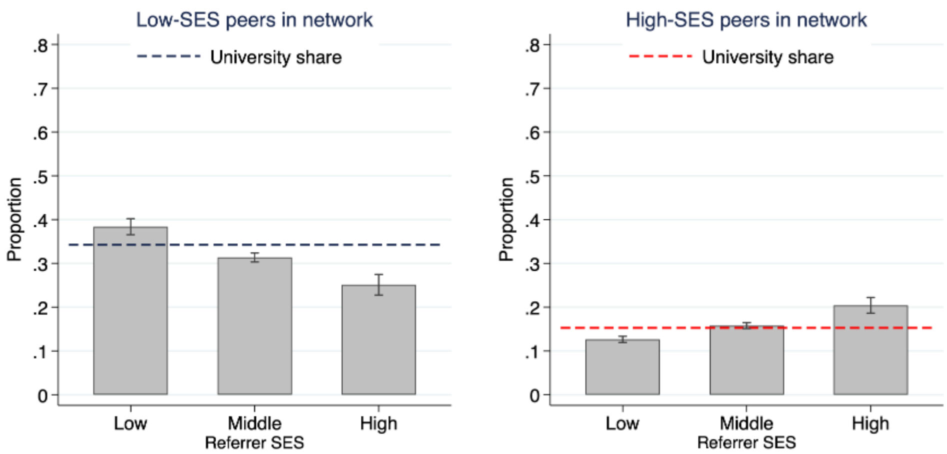
What are the diversity consequences of SES-driven differences across networks? In terms of network compositions, participants could connect with other SES groups at different rates than would occur randomly depending on their own SES. Figure 4.12a and Figure 4.12b illustrate the average network shares conditional on referrer SES respectively for low- and high-SES.[10] We observe modest deviations from university-wide SES shares in enrollment networks: Low-SES referrers have on average 38.4% low-SES peers compared to the university average of 34.3%, while high-SES referrers have 20.4% high-SES connections compared to the university average of 15.3%.

We find larger differences when studying connections between SES groups: Low-SES referrers connect with other low-SES at much higher rates than high-SES referrers (38.4% vs. 25.1%). Conversely, high-SES referrers connect more with other high-SES than low-

[10] For sake of brevity we omit middle-SES from this exposition. For the complete relationship, see Appendix Figure A.4.

SES referrers (20.4% vs. 12.6%). Middle-SES referrers are in between the two extreme patterns, connecting with middle-SES at higher rates than low-SES referrers (52.9% vs. 49.0%) but lower rates than high-SES referrers (52.9% vs. 54.5%). These findings indicate SES-based segregation in networks, with same-SES homophily across groups.

Figure 4.12: Network shares of SES groups

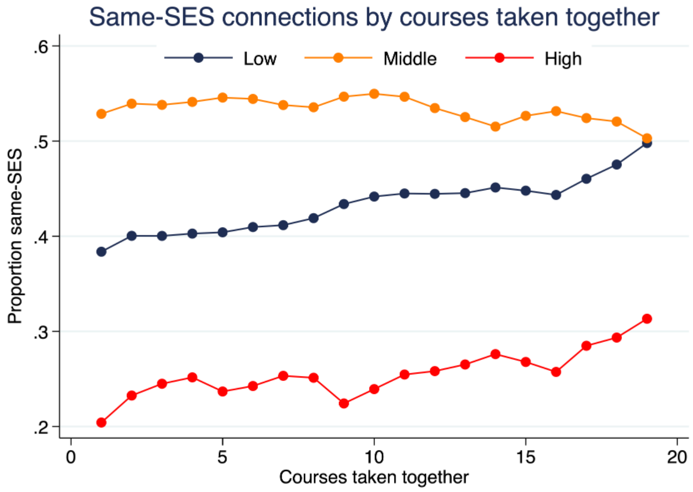


(a) Low-SES network share (b) High-SES network share

Note: Figures illustrate the average network shares of low- and high- SES peers conditional on referrer SES. Horizontal lines plot the university-wide shares of SES groups (low-SES: 34%, high-SES: 15%). While the share of low-SES peers in the network decreases as referrer SES increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

While same-SES students are connected more often with each other, so far we only consider the average the number of courses taken together with network members. What are the diversity implications of increased tie strength between students? As students take more courses together with peers, the share of same-SES peers in the networks of low- and high-SES increases while the share of middle-SES declines (see Figure 4.13). Both increases are substantial, amounting to 50% for high-, and 30% for low-SES beyond 15 courses together. While it is known that students who take courses together have similar characteristics (Kossinets & Watts, 2009), it is important to understand how increasing similarities in SES reflects on referral choice sets.

Figure 4.13: Network size and tie strength



Note: This figure illustrates the shares of same-SES connections for low-, middle-, and high-SES as a function of the average number of courses taken together with network members. Low- and high-SES networks both become more homogenous as the average number of courses taken together with their connections increase.

4.7.2 SES diversity in referral choice sets

How did the referrer choice sets look like in practice? We now combine our findings about network segregation with referral selection. In Section 4.6.2, we found that referrals went to peers with whom the median participant took 12 courses (average 14). By restricting the networks for courses taken above the median, we get an *ex post* snapshot of referrer choice sets.

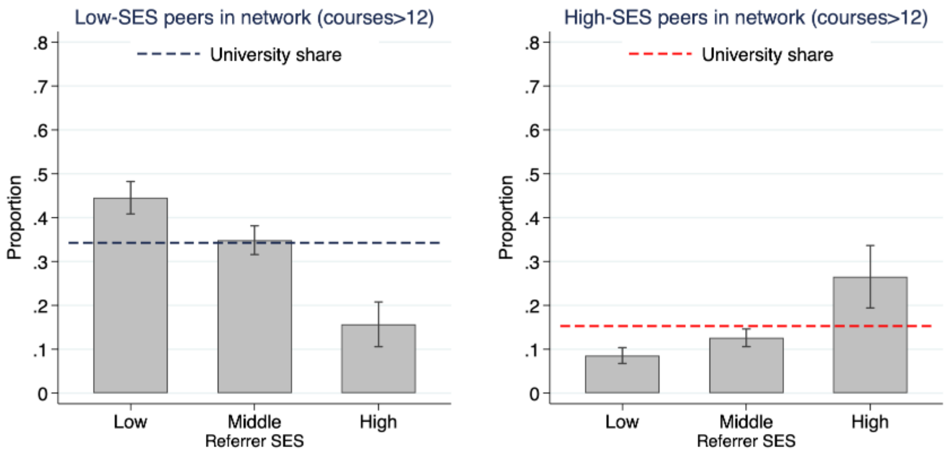
We show the average network shares conditional on referrer SES and above median number of courses taken together for low-SES in Figure 4.14a and for high-SES in Figure 4.14b.[11] Network compositions above the median number of courses taken reveal strong segregation effects in referral choice sets: Low-SES networks contain 44.5% low-SES peers, higher than the 34% university-wide share by 9.5 percentage points. Conversely, high-SES students are under-represented in low-SES networks at only 8.6% average share, compared to the 14% university share (−5.4 pp.). At the other extreme, high-SES networks show the reverse pattern with average low-SES share dropping to 15.7%, a 19.3

[11] In Appendix Figure A.5 we present the complete relationship including middle-SES.

percentage point decrease relative to the university average. High-SES students have a same-SES concentration at 26.5% (+12.5 pp.). Middle-SES networks remain relatively balanced and closely track university proportions.

Put differently, in an environment where 1 out of 3 students are low-SES, the chance that a low-SES peer is considered for a referral by high-SES is already less than 1/6. This stark disparity reveals that low-SES and high-SES students practically have separate networks within the same university, despite the opportunities to meet as equal-status students. The network segregation makes cross-SES referrals structurally unlikely even without any taste-based SES biases. We now explore program selection that emerges as a key driver of this segregation.

Figure 4.14: Network shares of SES groups above median tie strength



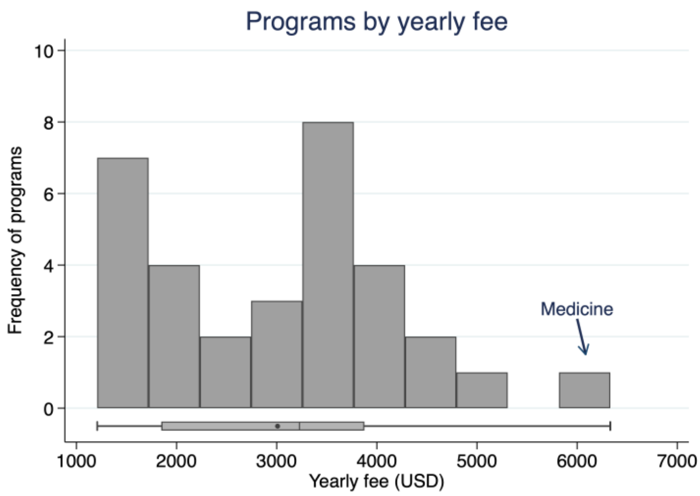
(a) Low-SES network share (b) High-SES network share

Note: Figures illustrate the average network shares of low- and high- SES peers conditional on referrer SES above the median number of courses taken together. Horizontal lines plot the university-wide shares of SES groups (low-SES: 34%, high-SES: 15%). While the share of low-SES peers in the network decreases as referrer SES increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

4.7.3 Program selection as a mechanism

Academic programs at this university have cost-based pricing, and typically less than 5% of students receive any kind of scholarship. Based on this, we first calculate how much every undergraduate program at the university is expected to cost students per year (see Figure 4.15). Considering that net minimum monthly wage stands at \$200 and the average Colombian salary around \$350, the cost differences between programs are large enough to make an impact on program selection. Is it the case that SES groups select into programs with financial considerations?

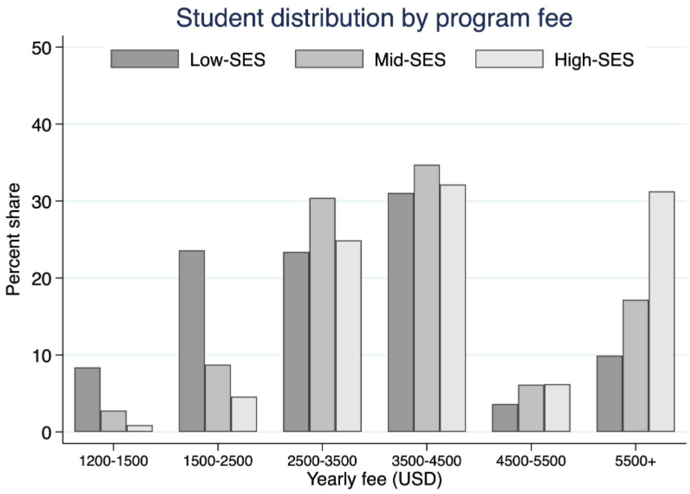
Figure 4.15: Undergraduate programs sorted by fee



Note: This figure illustrates the distribution of programs at the university by their average yearly fee. The average yearly fee stands at \$3000, and medicine is an outlier at \$6000.

To answer, we examine how SES groups are distributed across programs to identify evidence of SES-based selection (see Figure 4.16). Indeed, low-SES students select into more affordable programs, followed by middle-SES students. High-SES students sort almost exclusively into above-average costing programs, with a third selecting into medicine and creating a very skewed distribution. The distributions are significantly different across all pairwise comparisons: low-SES vs. middle-SES (Kolmogorov-Smirnov test $D = 33.89$, $p < 0.001$), low-SES vs. high-SES ($D = 31.31$, $p < 0.001$), and middle-SES vs. high-SES ($D = 31.31$, $p < 0.001$). These findings support the idea that program selection could be the reason why low- and high-SES networks tend to segregate as the number of courses taken increases. Financial constraints channel students into different academic programs, which in turn determine their classroom interactions and university social networks.

Figure 4.16: SES distribution by program fee



Note: This figure illustrates the distribution of each SES group across programs sorted by fee. The majority of low-SES select into programs with below average cost, while high-SES select into programs with above average cost. Medicine accounts for a third of all high-SES students at this university.

4.7.4 Robustness check: Tie strength and sharing academic programs

Does the number of courses taken together have an independent effect that goes beyond identifying peers in the same academic program? To evaluate this question we leverage our administrative data, and identify peers within the same program: In each individual network we observe the participant-specific academic program for the referrer and alternative-specific academic program for each referral candidate. We add this new variable in our specification and describe our findings in Table 4.7. Being in the same academic program has a substantial positive effect on referral likelihood, with coefficients ranging from 1.257 to 2.198 across all referrer SES groups. This confirms that program affiliation serves as a strong predictor of referral decisions. Our comparison of interest is the point estimate for the standardized number of courses taken. Across all three referrer SES groups, the standardized number of courses taken together maintains its statistical significance after controlling for same program membership. The coefficient magnitudes are expectedly smaller compared to specifications without program controls (ranging from 0.688 to 0.930) as the newly added variable is a moderator: Matching academic programs leads to taking more courses together. The remaining estimates in our model are robust to the inclusion of the same-program variable with little change in point estimates. The persistence of statistical significance (all $p < 0.001$) suggests that the number of courses taken together has an independent effect on referral decisions. To sum, our measure of tie strength seems to capture meaningful social interaction patterns that lead to referrals, and go beyond simply identifying matching academic programs.

Table 4.7: SES bias in referral decisions by referrer SES group with program controls

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.236**	-0.140	-0.567*
	(0.119)	(0.111)	(0.331)
High-SES referral	-0.421*	-0.249	-0.383
	(0.220)	(0.158)	(0.281)
Entry exam (referral z-score)	0.623***	0.590***	0.892***
	(0.054)	(0.048)	(0.114)
Courses taken (z-score)	0.688***	0.760***	0.930***
	(0.032)	(0.035)	(0.119)
Same program	2.074***	2.198***	1.257***
	(0.215)	(0.185)	(0.467)
χ^2	865.35	981.99	135.47
Observations	110,142	127,088	19,767
Individuals	301	366	67

Note: Individual-level clustered standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a separate conditional logit regression estimated on the subsample of referrers from the indicated SES group. Coefficients represent log-odds of referring candidates from the specified SES group relative to referring middle-SES candidates. All models include individual fixed effects that control for each referrer's choice set composition.

4.8 Conclusion

We investigate whether SES biases in referral selection stem from individual preferences in choosing an SES group over others or network segregation. Through a lab-in-the-field experiment with 734 university students making incentivized referrals from their enrollment networks, we find that institutional factors dominate individual preferences.

Our key findings are threefold. First, referral patterns remain unchanged across different incentive structures: participants consistently select high-performing peers with a high number of courses taken together regardless of whether referral recipients receive additional compensation. Second, we find an SES bias is that is asymmetric and limited. While low-SES referrers exhibit strong in-group preferences, middle- and high-SES referrers show no bias toward their own and other groups. Third, network segregation driven by cost-based program selection explains most referral patterns. At typical referral range measured by the number of courses taken together, low-SES and high-SES students have dramatically different choice sets, with high-SES networks containing only 15.7% low-SES peers compared to 34% university-wide.

These results have important policy implications. While universities expose low-SES students to higher-than-population shares of high-SES peers, segregation within institutions limits meaningful interaction across SES. Our findings suggest that institutional interventions promoting cross-SES contact, represents a promising approach to reduce SES-based inequality in opportunity transmission. Future research should explore the causal effects of specific institutional interventions such as mixed seating (Rohrer et al., 2021), or cross-SES mentoring programs (Alan & Kubilay, 2025), that increase interactions between SES groups.

References

Alan, S., & Kubilay, E. (2025). Empowering Adolescents to Transform Schools: Lessons from a Behavioral Targeting. *American Economic Review*, *115*(2), 365–407. doi: 10.1257/aer.20240374

Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en colombia. *Documentos CEDE*.

Beaman, L., Keleher, N., & Magruder, J. (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics*, *36*(1), 121–157. doi: 10.1086/693869

Beaman, L., & Magruder, J. (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review*, *102*(7), 3574–3593. doi: 10.1257/aer.102.7.3574

Bolte, L., Jackson, M. O., & Immorlica, N. (2024). The Role of Referrals in Immobility, Inequality, and Inefficiency in Labor Markets. *Journal of Labor Economics*. doi: 10.1086/733048

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood Press.

Brown, M., Setren, E., & Topa, G. (2016). Do informal referrals lead to better matches? evidence from a firm's employee referral system. *Journal of Labor Economics*, *34*(1), 161–209.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... Wernerfelt, N. (2022a). Social capital 1: Measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121. doi: 10.1038/s41586-022-04996-4

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... Wernerfelt, N. (2022b). Social capital 2: Determinants of economic connectedness. *Nature*, *608*(7921), 122–134. doi: 10.1038/s41586-022-04997-3

Currarini, S., Jackson, M. O., & Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, *77*(4), 1003–1045. doi: 10.3982/ECTA7528

Currarini, S., Jackson, M. O., & Pin, P. (2010). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National*

Academy of Sciences, *107*(11), 4857–4861. doi: 10.1073/pnas.0911793107

Dustmann, C., Glitz, A., Schönberg, U., & Brücker, H. (2016). Referral-based job search networks. *The Review of Economic Studies*, *83*(2), 514–546.

Engzell, P., & Wilmers, N. (2025). Firms and the Intergenerational Transmission of Labor Market Advantage. *American Journal of Sociology*, 736993. doi: 10.1086/736993

Fergusson, L., & Flórez, S. A. (2021a). Desigualdad educativa en colombia. In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel.

Fergusson, L., & Flórez, S. A. (2021b). Distinción escolar. In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel.

Friebel, G., Heinz, M., Hoffman, M., & Zubanov, N. (2023). What do employee referral programs do? measuring the direct and overall effects of a management practice. *Journal of Political Economy*, *131*(3), 633–686.

Friedman, S., & Laurison, D. (2019). Getting on. In *The class ceiling: Why it pays to be privileged* (pp. 45–56). Bristol, UK and Chicago, IL, USA: Policy Press. (Chapter 2 of the authored book.)

Galos, D. R. (2024). Social media and hiring: A survey experiment on discrimination based on online social class cues. *European Sociological Review*, *40*(1), 116–128. doi: 10.1093/esr/jcad012

García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la cuna: La movilidad social a través de la educación en los municipios de colombia. *Documentos CEDE*.

García Villegas, M., & Cobo, P. (2021). La dimensión cultural del apartheid educativo. In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel.

Gee, L. K., Jones, J., & Burke, M. (2017). Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network. *Journal of Labor Economics*, *35*(2), 485–518. doi: 10.1086/686225

Griffith, A. (2022). Name Your Friends, but Only Five? The Importance of Censoring in

Peer Effects Estimates Using Social Network Data. *Journal of Labor Economics*. doi: 10.1086/717935

Guevara S, J. D., & Shields, R. (2019). Spatializing stratification: Bogotá. *Ardeth. A Magazine on the Power of the Project*(4), 223–236.

Hederos, K., Sandberg, A., Kvissberg, L., & Polano, E. (2025). Gender homophily in job referrals: Evidence from a field study among university students. *Labour Economics*, *92*, 102662.

Hudson, R. A., & Library of Congress (Eds.). (2010). *Colombia: a country study* (5th ed.). Washington, D.C: Federal Research Division, Library of Congress: For sale by the Supt. of Docs., U.S. G.P.O. Retrieved from the Library of Congress, https://www.loc.gov/item/2010009203/.

Jackson, M. O. (2022). *Inequality's Economic and Social Roots: The Role of Social Networks and Homophily* (SSRN Scholarly Paper No. 3795626). Rochester, NY. doi: 10.2139/ssrn.3795626

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, *115*(2), 405–450. Retrieved from https://www.journals.uchicago.edu/doi/abs/10.1086/599247 doi: 10.1086/599247

Kramarz, F., & Skans, O. N. (2014). When strong ties are strong: Networks and youth labour market entry. *Review of economic studies*, *81*(3), 1164–1200.

Laurison, D., & Friedman, S. (2024). The Class Ceiling in the United States: Class-Origin Pay Penalties in Higher Professional and Managerial Occupations. *Social Forces*, *103*(1), 22–44. doi: 10.1093/sf/soae025

Lin, N., Ensel, W. M., & Vaughn, J. C. (1981). Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment. *American Sociological Review*, *46*(4), 393–405. doi: 10.2307/2095260

Loury, G. C. (1977). A dynamic theory of racial income differences. In P. A. Wallace & A. M. LaMond (Eds.), *Women, minorities, and employment discrimination* (pp. 153–186). Lexington, MA: Lexington Books. (Originally published as Discussion Paper 225, Northwestern University, Center for Mathematical Studies in Economics and Management Science, 1976)

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, *27*(1), 415–444.

Michelman, V., Price, J., & Zimmerman, S. D. (2022). Old Boys' Clubs and Upward Mobility Among the Educational Elite. *The Quarterly Journal of Economics*, *137*(2), 845–909. doi: 10.1093/qje/qjab047

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *American Economic Review*.

Mouw, T. (2003). Social Capital and Finding a Job: Do Contacts Matter? *American Sociological Review*, *68*(6), 868–898. doi: 10.1177/000312240306800604

Núñez, J., & Gutiérrez, R. (2004). Class discrimination and meritocracy in the labor market: evidence from chile. *Estudios de Economía*, *31*(2), 113–132.

Obukhova, E., & Lan, G. (2013). Do Job Seekers Benefit from Contacts? A Direct Test with Contemporaneous Searches. *Management Science*, *59*(10), 2204–2216. doi: 10.1287/mnsc.1120.1701

Pallais, A., & Sands, E. G. (2016). Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*, *124*(6), 1793–1828. doi: 10.1086/688850

Pedulla, D. S., & Pager, D. (2019). Race and networks in the job search process. *American Sociological Review*, *84*, 983-1012. doi: 10.1177/0003122419883255

Rivera, L. A. (2012). Hiring as Cultural Matching: The Case of Elite Professional Service Firms. *American Sociological Review*, *77*(6), 999–1022. doi: 10.1177/0003122412463213

Rivera, L. A., & Tilcsik, A. (2016). Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market. *American Sociological Review*, *81*(6), 1097–1131. doi: 10.1177/0003122416668154

Rohrer, J. M., Keller, T., & Elwert, F. (2021). Proximity can induce diverse friendships: A large randomized classroom experiment. *PLOS ONE*, *16*(8), e0255097. doi: 10.1371/journal.pone.0255097

Smith, S. S. (2005). "Don't put my name on it": Social Capital Activation and Job-Finding Assistance among the Black Urban Poor. *American Journal of Sociology*, *111*(1), 1–57. doi: 10.1086/428814

Stansbury, A., & Rodriguez, K. (2024). The class gap in career progression: Evidence from US academia. *Working Paper*.

Sterling, A. D. (2014). Friendships and Search Behavior in Labor Markets. *Management*

Science, *60*(9), 2341–2354. doi: 10.1287/mnsc.2013.1857

Topa, G. (2019). Social and spatial networks in labour markets. *Oxford Review of Economic Policy*, *35*(4), 722–745.

United Nations. (2023). *Social panorama of latin america and the caribbean 2023: labour inclusion as a key axis of inclusive social development.* ECLAC and United Nations. Retrieved from https://www.cepal.org/es/publicaciones/68702-panorama-social-america-latina-caribe-2023-la-inclusion-laboral-como-eje-central

Uribe-Mallarino, C. (2008). Estratificación social en bogotá: de la política pública a la dinámica de la segregación social. *Universitas humanistica*(65), 139–172.

Wang, S.-Y. (2013). Marriage networks, nepotism, and labor market outcomes in China. *American Economic Journal: Applied Economics*, *5*(3), 91–112.

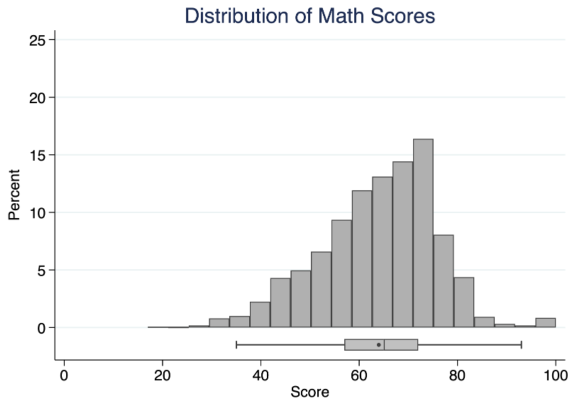
Witte, M. (2021). Why do workers make job referrals? experimental evidence from ethiopia. *Working Paper*.

World Bank. (2024). *Regional poverty and inequality update spring 2024* (Poverty and Equity Global Practice Brief). Washington, D.C.: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269

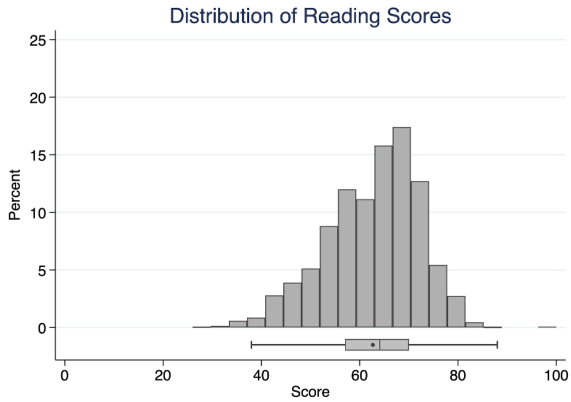
A.1 Additional Figures and Tables

Additional Figures

Figure A.1: Distribution of exam scores at the university



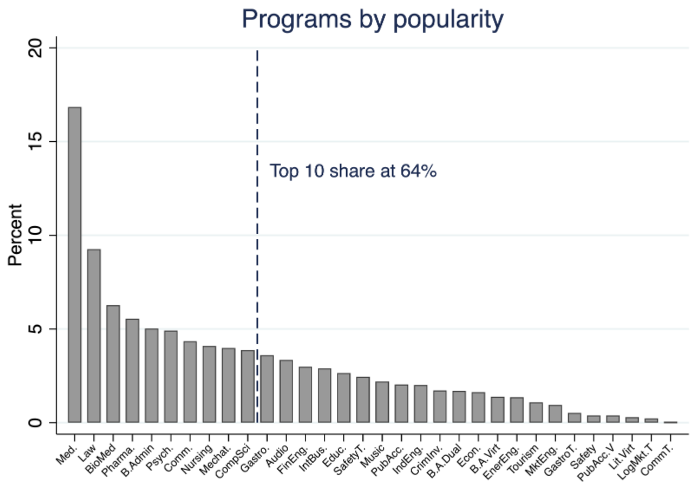
(a) Math scores at the university



(b) Reading scores at the university

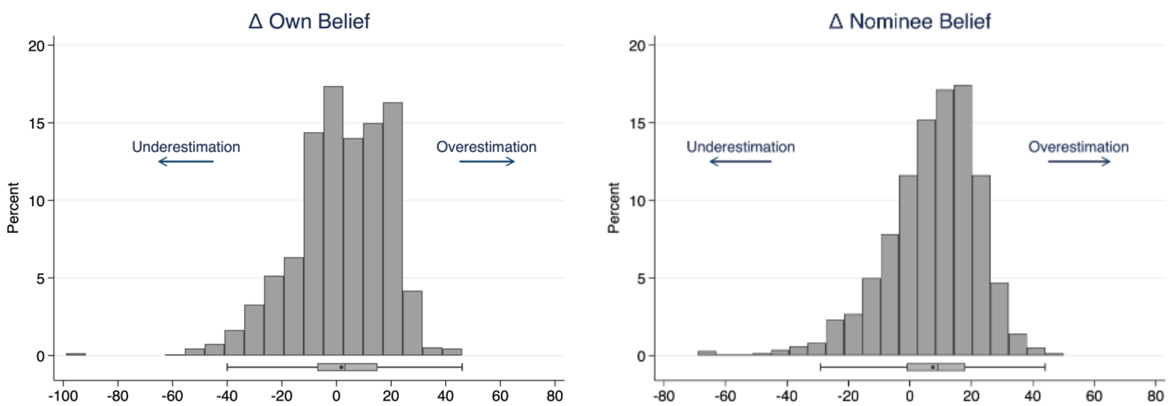
Note: Reading scores (left panel) and math scores (right panel) show tight distributions with approximately 75% of students falling within just 13-15 points of each other.

Figure A.2: Distribution of students across undergraduate programs



Note: This figure shows the concentration of students across 32 undergraduate programs at the university. Students cluster around certain programs. The top 5 most popular programs (Medicine, Law, Biomedical Engineering, Pharmacy Technology, and Business Administration) account for 43% of all undergraduates, and the top 10 most popular programs account for 64% of students.

Figure A.3: Distribution of participant beliefs

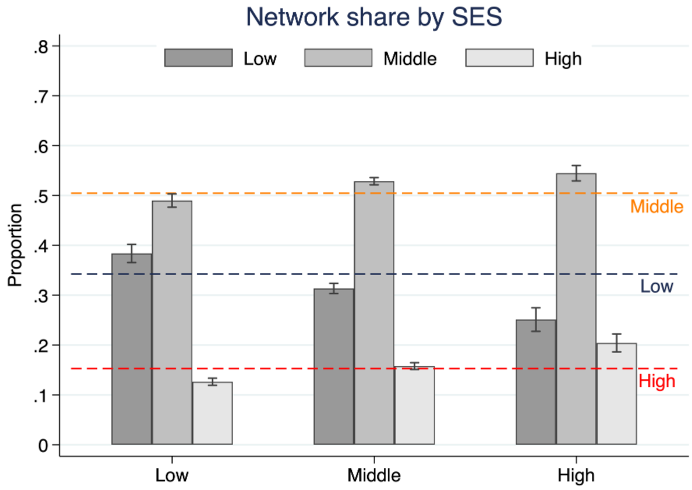


(a) Beliefs about own performance rank

(b) Beliefs about referral performance rank

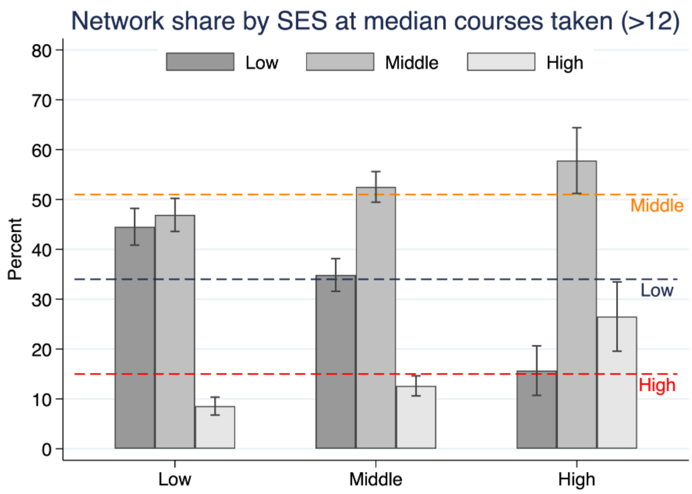
Note: Panel (a) illustrates participant's own rank belief minus their actual rank at the university. Panel (b) illustrates participant's rank belief of their referral minus their actual rank at the university. While participants accurately assess their own rank, they slightly overestimate their referral's rank.

Figure A.4: Network shares by SES



Note: This figure displays the average network shares of SES groups respectively for low-, middle-, and high-SES referrers. Horizontal lines plot the university-wide shares of each SES group (Low: 34%, Mid: 51%, High: 15%). While the share of low-SES peers in the network decreases as the SES of the referrers increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

Figure A.5: Network shares by SES at courses taken above 12



Note: This figure displays the average network shares of SES groups respectively for low-, middle-, and high-SES referrers above the median number of courses taken together. Horizontal lines plot the university-wide shares of each SES group (Low: 34%, Mid: 51%, High: 15%). Low- and high-SES networks both become same-SES dominated at the expense of each other while middle-SES networks remain balanced. Error bars represent 95% confidence intervals.

Additional Tables

Table A.1: Selection into the experiment

	University	**Sample**	***p***
Reading score	62.651	65.183	< 0.001
Math score	63.973	67.477	< 0.001
GPA	3.958	4.012	< 0.001
Low-SES	0.343	0.410	< 0.001
Middle-SES	0.505	0.499	0.763
High-SES	0.153	0.091	< 0.001
Female	0.567	0.530	< 0.001
Age	21.154	20.651	< 0.001
Observations	4,417	734	

Note: This table compares characteristics between the university and the experimental sample. p-values for binary outcomes (Low-SES, Med-SES, High-SES, Female) are from two-sample tests of proportions; for continuous variables, from two-sample t-tests with unequal variances. All reported p-values are two-tailed.

Table A.2: Distribution of referrals by area

Area	**Only one area**	**Both areas**	**Total**
Verbal	65	608	673
Math	61	608	669
Total	126	1,216	1,342

Note: The table shows how many referrers made referrals in only one area versus both areas. "Only one area" indicates individuals who made referrals exclusively for one area of the exam. "Both areas" shows individuals who made referrals in both verbal and math areas. The majority of referrers (608) made referrals in both areas.

Table A.3: Referral characteristics by exam area (unique referrals only)

	Reading	**Math**	***p***
Reading score	67.733	67.126	0.252
Math score	69.339	71.151	0.008
GPA	4.136	4.136	0.987
Courses taken	13.916	13.019	0.123
Low-SES	0.372	0.385	0.666
Med-SES	0.526	0.518	0.801
High-SES	0.103	0.097	0.781
Observations	487	483	

Note: This table compares characteristics of uniquely referred students by entry exam area for the referral (verbal vs. math). p-values are from two-sample t-tests with unequal variances. Referrals in Math area go to peers with significantly higher math scores ($p = 0.008$), while we find no significant differences for Reading scores, GPA, courses taken, or SES composition for referrals across the two areas. Excluding referrals going to the same individuals does not change the outcomes for referrals compared to Appendix Table A.4

Table A.4: Referral characteristics by academic area

	Reading	**Math**	***p***
Reading score	67.85	67.41	0.348
Math score	70.04	71.36	0.029
GPA	4.153	4.153	0.984
Courses taken	14.467	13.822	0.206
Low-SES	37%	38%	0.714
Middle-SES	51%	51%	0.829
High-SES	11%	11%	0.824
Observations	673	669	

Note: This table compares characteristics of referred students by entry exam area for the referral (verbal vs. math). p-values are from two-sample t-tests with unequal variances. Referrals in Math area go to peers with significantly higher math scores ($p = 0.029$), while we find no significant differences for Reading scores, GPA, courses taken, or SES composition for referrals across the two areas.

Table A.5: Average entry exam z-scores by SES network connections

	Network average for SES group		
Referrer SES	**Low**	**Middle**	**High**
Low	0.086	-0.018	0.144
Middle	0.186	0.023	0.215
High	0.204	0.064	0.285
All	-0.361	-0.078	0.169

Note: This table shows average (math and critical reading) standardized entry exam scores for individuals of different SES levels (rows) when connected to peers of specific SES levels (columns). The "All" row shows the overall average scores across all participant SES levels when connected to each network SES type. Higher values indicate better academic performance in SD's.

A.2 Experiment

We include the English version of the instructions used in Qualtrics. Participansts saw the Spanish version. Horizontal lines in the text indicate page breaks and clarifiying comments are inside brackets.

Consent

You have been invited to participate in this decision-making study. This study is directed by [omitted for anonymous review] and organized with the support of the Social Bee Lab (Social Behavior and Experimental Economics Laboratory) at UNAB.

In this study, we will pay **one (1)** out of every **ten (10)** participants, who will be randomly selected. Each selected person will receive a fixed payment of **70,000** (seventy thousand pesos) for completing the study. Additionally, they can earn up to **270,000** (two hundred and seventy thousand pesos), depending on their decisions. So, in total, if you are selected to receive payment, you can earn up to **340,000** (three hundred and forty thousand pesos) for completing this study.

If you are selected, you can claim your payment at any Banco de Bogotá office by presenting your ID. Your participation in this study is voluntary and you can leave the study at any time. If you withdraw before completing the study, you will not receive any payment.

The estimated duration of this study is 20 minutes.

The purpose of this study is to understand how people make decisions. For this, we will use administrative information from the university such as the SABER 11 test scores of various students (including you). Your responses will not be shared with anyone and your participation will not affect your academic records. To maintain strict confidentiality, the research results will not be associated at any time with information that could personally identify you.

There are no risks associated with your participation in this study beyond everyday risks. However, if you wish to report any problems, you can contact Professor [omitted for anonymous review]. For questions related to your rights as a research study participant, you can contact the IRB office of [omitted for anonymous review].

By selecting the option "I want to participate in the study" below, you give your consent to participate in this study and allow us to compare your responses with some administrative records from the university.

- I want to participate in the study [advances to next page]
- I do not want to participate in the study

Student Information

Please write your student code. In case you are enrolled in more than one program simultaneously, write the code of the first program you entered:

[Student ID code]

What semester are you currently in?

[Slider ranging from 1 to 11]

[Random assignment to treatment or control]

Instructions

The instructions for this study are presented in the following video. Please watch it carefully. We will explain your participation and how earnings are determined if you are selected to receive payment.

[Treatment-specific instructions in video format]

If you want to read the text of the instructions narrated in the video, press the "Read instruction text" button. Also know that in each question, there will be a button with information that will remind you if that question has earnings and how it is calculated, in case you have any doubts.

- I want to read the instructions text [text version below]

In this study, you will respond to three types of questions. First, are the belief questions. For belief questions, we will use as reference the results of the SABER 11 test that you and other students took to enter the university, focused on three areas of the exam: mathematics, reading, and English.

For each area, we will take the scores of all university students and order them from lowest to highest. We will then group them into 100 percentiles. The percentile is a position measure that indicates the percentage of students with an exam score that is above or below a value.

For example, if your score in mathematics is in the 20th percentile, it means that 20 percent of university students have a score lower than yours and the remaining 80 percent have a higher score. A sample belief question is: "compared to university students, in what percentile is your score for mathematics?"

If your answer is correct, you can earn 20 thousand pesos. We say your answer is correct if the difference between the percentile you suggest and the actual percentile of your score is not greater than 7 units. For example, if you have a score that is in the 33rd percentile and you say it is in the 38th, the answer is correct because the difference is less than 7. But if you answer that it is in the 41st, the difference is greater than 7 and the answer is incorrect.

The second type of questions are recommendation questions and are also based on the mathematics, reading, and English areas of the SABER 11 test. We will ask you to think about the students with whom you have taken or are taking classes, to recommend from among them the person you consider best at solving problems similar to those on the SABER 11 test.

When you start typing the name of your recommended person, the computer will show suggestions with the full name, program, and university entry year of different students. Choose the person you want to recommend. If the name doesn't appear, check that you are writing it correctly. Do not use accents and use 'n' instead of 'ñ'. If it still doesn't appear, it may be because that person is not enrolled this semester or because they did not take the SABER 11 test. In that case, recommend someone else.

You can earn up to 250,000 pesos for your recommendation. We will multiply your recommended person's score by 100 pesos if they are in the first 50 percentiles. We will multiply it by 500 pesos if your recommended person's score is between the 51st and 65th percentile. If it is between the 66th and 80th percentile, we will multiply your recommended person's score by 1000 pesos. If the score is between the 81st and 90th percentile, you earn 1500 pesos multiplied by your recommended person's score. And if the score is between the 91st and 100th percentile, we will multiply your recommended person's score by 2500 pesos to determine the earnings.

The third type of questions are information questions and focus on aspects of your personal life or your relationship with the people you have recommended.

Earnings

Now we will explain who gets paid for participating and how the earnings for this study are assigned. The computer will randomly select one out of every 10 participants to pay for their responses. For selected individuals, the computer will randomly choose one of the three areas, and from that chosen area, it will pay for one of the belief questions.

Similarly, the computer will randomly select one of the three areas to pay for one of the recommendation questions.

Additionally, if you are selected to receive payment, your recommended person in the chosen area will receive a fixed payment of 100 thousand pesos. [Only seen if assigned to the treatment]

Each person selected to receive payment for this study can earn: up to 20 thousand pesos for one of the belief questions, up to 250 thousand pesos for one of the recommendation questions, and a fixed payment of 70 thousand pesos for completing the study.

Selected individuals can earn up to 340 thousand pesos.

[Participants go through all three Subject Areas in randomized order]

Subject Areas

Critical Reading

For this section, we will use as reference the Critical Reading test from SABER 11, which evaluates the necessary competencies to understand, interpret, and evaluate texts that can be found in everyday life and in non-specialized academic fields.

[Clicking shows the example question from SABER 11 below]

Although the democratic political tradition dates back to ancient Greece, political thinkers did not address the democratic cause until the 19th century. Until then, democracy had been rejected as the government of the ignorant and unenlightened masses. Today it seems that we have all become democrats without having solid arguments in favor. Liberals, conservatives, socialists, communists, anarchists, and even fascists have rushed to proclaim the virtues of democracy and to show their democratic credentials (Andrew Heywood). According to the text, which political positions identify themselves as democratic?

- Only political positions that are not extremist
- The most recent political positions historically
- The majority of existing political positions
- The totality of possible political currents

Mathematics

This section references the Mathematics test from SABER 11, which evaluates people's competencies to face situations that can be resolved using certain mathematical tools.

[Clicking shows the example question from SABER 11 below]

A person living in Colombia has investments in dollars in the United States and knows that the exchange rate of the dollar against the Colombian peso will remain constant this month, with 1 dollar equivalent to 2,000 Colombian pesos. Their investment, in dollars, will yield profits of 3% in the same period. A friend assures them that their profits in pesos will also be 3%. Their friend's statement is:

- Correct. The proportion in which the investment increases in dollars is the same as in pesos.
- Incorrect. The exact value of the investment should be known.
- Correct. 3% is a fixed proportion in either currency.
- Incorrect. 3% is a larger increase in Colombian pesos.

English

This section uses the English test from SABER 11 as a reference, which evaluates that the person demonstrates their communicative abilities in reading and language use in this language.

[Clicking shows the example question from SABER 11 below]

Complete the conversations by marking the correct option.

- Conversation 1: I can't eat a cold sandwich. It is horrible!
  - I hope so.
  - I agree.
  - I am not.
- Conversation 2: It rained a lot last night!
  - Did you accept?
  - Did you understand?
  - Did you sleep?

[Following parts are identical for all Subject Areas and are not repeated here for brevity]

Your Score

Compared to university students, in which percentile do you think your [**Subject Area**] test score falls (1 is the lowest percentile and 100 the highest)?

[Clicking shows the explanations below]

How is a percentile calculated?

A percentile is a position measurement. To calculate it, we take the test scores for all students currently enrolled in the university and order them from lowest to highest. The percentile value you choose refers to the percentage of students whose score is below yours. For example, if you choose the 20th percentile, you're indicating that 20% of students have a score lower than yours and the remaining 80% have a score higher than yours.

What can I earn for this question?

For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the difference between your response and the correct percentile is less than 7. For example, if the percentile where your score falls is 33 and you respond with 38 (or 28), the difference is 5 and the answer is considered correct. But if you respond with 41 or more (or 25 or less), for example, the difference would be greater than 7 and the answer is incorrect.

Please move the sphere to indicate which percentile you think your score falls in:

[Slider with values from 0 to 100]

Recommendation

Among the people with whom you have taken any class at the university, who is your recommendation for the [**Subject Area**] test? Please write that person's name in the box below:

Important: You will not be considered for payment unless the recommended person is someone with whom you have taken at least one class during your studies.

Your response is only a recommendation for the purposes of this study and we will **not** contact your recommended person at any time.

[Clicking shows the explanations below]

Who can I recommend?

Your recommendation **must** be someone with whom you have taken (or are taking) a class. If not, your answer will not be considered for payment. **The person you recommend will not be contacted or receive any benefit from your recommendation.** [Only seen if assigned to the treatment]

As you write, you will see up to 7 suggested student names containing the letters you have entered. The more you write, the more accurate the suggestions will be. Please write **without** accents and use the letter 'n' instead of 'ñ'. If the name of the person you're writing doesn't appear, it could be because you made an error while writing the name.

If the name is correct and still doesn't appear, it could be because the student is not enrolled this semester or didn't take the SABER 11 test. In that case, you must recommend someone else.

My earnings for this question?

For your recommendation, you could receive earnings of up to 250,000 (two hundred and fifty thousand) PESOS. The earnings are calculated based on your recommendation's score and the percentile of that score compared to other UNAB students, as follows:

- We will multiply your recommendation's score by $100 (one hundred) pesos if it's between the 1st and 50th percentiles
- We will multiply your recommendation's score by $500 (five hundred) pesos if it's between the 51st and 65th percentiles
- We will multiply your recommendation's score by $1000 (one thousand) pesos if it's between the 66th and 80th percentiles
- We will multiply your recommendation's score by $1500 (one thousand five hundred) pesos if it's between the 81st and 90th percentiles
- We will multiply your recommendation's score by $2500 (two thousand five hundred) pesos if it's between the 91st and 100th percentiles

This is illustrated in the image below:

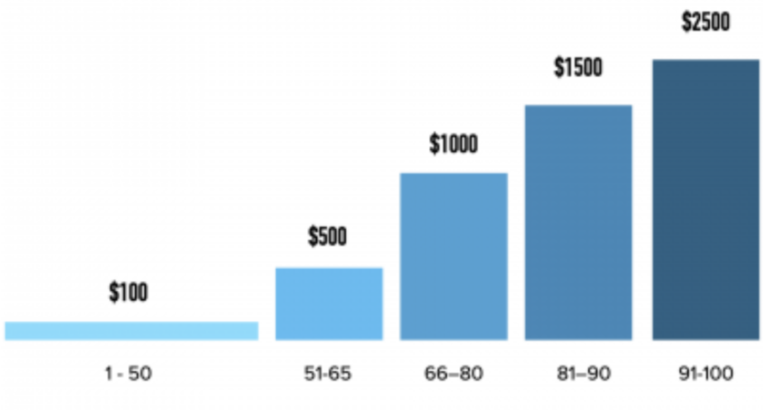


Figure B.1: Earnings for recommendation questions

For example, if your recommendation got 54 points and the score is in the 48th percentile, you could earn 54x100 = 5400 PESOS. But, if the same score of 54 points were in the 98th percentile, you could earn 54x2500 = 135,000 PESOS.

[Text field with student name suggestions popping up as participant types]

Relationship with your recommendation

How close is your relationship with your recommendedation: "[Name of the student selected from earlier]"? (0 indicates you are barely acquaintances and 10 means you are very close)

[Slider with values from 0 to 10]

Your recommendation's score

Compared to university students, in which percentile do you think [Name of the student selected from earlier]'s score falls in the [**Subject Area**] test (1 is the lowest percentile and 100 the highest)?

[Clicking shows the explanations below]

How is a percentile calculated?

A percentile is a position measurement. To calculate it, we take the test scores for all students currently enrolled in the university and order them from lowest to highest. The percentile value you choose refers to the percentage of students whose score is below yours. For example, if you choose the 20th percentile, you're indicating that 20% of students have a score lower than yours and the remaining 80% have a score higher than yours.

What can I earn for this question?

For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the difference between your response and the correct percentile is less than 7. For example, if the percentile where your recommended person's score falls is 33 and you respond with 38 (or 28), the difference is 5 and the answer is considered correct. But if you respond with 41 or more (or 25 or less), for example, the difference would be greater than 7 and the answer is incorrect.

Please move the sphere to indicate which percentile you think your recommended person's score falls in:

[Slider with values from 0 to 100]

Demographic Information

What is the highest level of education achieved by your father?

[Primary, High School, University, Graduate Studies, Not Applicable]

What is the highest level of education achieved by your mother?

[Primary, High School, University, Graduate Studies, Not Applicable]

Please indicate the socio-economic group to which your family belongs:

[Group A (Strata 1 or 2), Group B (Strata 3 or 4), Group C (Strata 5 or 6)]

UNAB Students Distribution

Thinking about UNAB students, in your opinion, what percentage belongs to each socio-economic group? The total must sum to 100%:

[Group A (Strata 1 or 2) percentage input area]

[Group B (Strata 3 or 4) percentage input area]

[Group C (Strata 5 or 6) percentage input area]

[Shows sum of above percentages]

End of the Experiment

Thank you for participating in this study.

If you are chosen to receive payment for your participation, you will receive a confirmation to your UNAB email and a link to fill out a form with your information. The process of processing payments is done through Nequi and takes approximately 15 business days, counted from the day of your participation.

[Clicking shows the explanations below]

Who gets paid and how is it decided?

The computer will randomly select one out of every ten participants in this study to be paid for their decisions.

For selected individuals, the computer will randomly select one area: mathematics, reading, or English, and from that area will select one of the belief questions. If the answer to that question is correct, the participant will receive 20,000 pesos.

The computer will randomly select an area (mathematics, critical reading, or English) to pay for one of the recommendation questions. The area chosen for the recommendation question is independent of the area chosen for the belief question. The computer will take one of the two recommendations you have made for the chosen area. Depending on your recommendation's score, you could win up to 250,000 pesos.

Additionally, people selected to receive payment for their participation will have a fixed earnings of 70,000 pesos for completing the study.

Participation

In the future, we will conduct studies similar to this one where people can earn money for their participation. The participation in these studies is by invitation only. Please indicate if you are interested in being invited to other studies similar to this one:

[Yes, No]

Chapter 5

Conclusion

This dissertation set out to examine economic behavior under environmental and structural constraints, motivated by the wide body of evidence that people diverge from the all-knowing and all-capable rationality of *homo economicus*. Through three essays exploring the effects of digital nudging and social networks, I have attempted to document several ways in which outside factors shape economic decisions beyond individual preferences.

My first chapter sought to isolate the causal effects of digital design features on consumer behavior. The investigation into autoplay, however, produced findings that contradicted my hypotheses and revealed the challenges of studying real-world behavior in a laboratory setting. For instance, autoplay did not increase video consumption, yet participants were willing to pay to keep it active. These puzzling results stemmed directly from the experimental design, which influenced participant behavior in ways that confounded the effects of my treatment. This experience, combined with intertemporal choice literature showing that the framing of tasks can produce contentious results, proved invaluable and highlighted the importance of context. It became clear that to understand a feature like autoplay, research must account for the personalized, algorithmically-driven environments where it is actually deployed. This realization guided the shift toward more naturalistic settings in the subsequent chapters.

The second and third chapters, examined how institutional constraints shape the transmission of economic opportunities through social networks. The findings from these chapters paint a nuanced picture of how cross-SES contact takes place in higher education. Chapter two, for example, demonstrated that sustained classroom interaction can be effective at reducing bias and help students learn about their peers' cognitive skill and academic performance. This semester-long interaction between classmates resulted in successful referrals of higher performing peers inside the classroom, and targeted incentives remedied the small SES biases that existed.

However, chapter three revealed the sobering reality that such positive classroom interactions represent only a small fraction of students' university experience. The institutional structure of the university, particularly program selection driven by the fees of these programs, creates powerful forces that segregate the networks of social groups. As low- and high-SES students progress through their studies, taking more courses within their programs, their networks become increasingly same-SES homogeneous. The result is that despite the university's apparent diversity and the demonstrated potential for positive cross-class interactions, meaningful connections between high- and low-SES students remain limited.

These findings carry implications for institutional design. In the Colombian higher education context, non-elite private universities do provide returns to low-SES in terms of social capital, but it depends on the frequency of cross-SES interactions inside the classroom. Institutional decisions have profound impacts on the extent of social capital accumulation. For example, policies like the cost-based program fee structure that ends up sorting students by income undermines the potential benefits from housing a very diverse student body.

Several broader conclusions emerge from this research. First, setting up the appropriate experimental context matters enormously for gathering policy relevant evidence. Digital nudges created by private firms and the efforts to regulate these point to the fact that individuals who were not tempted by autoplay in the laboratory may be susceptible to its effects in their daily lives. If the objective is to guide policymakers, researchers need concrete evidence from the environment where the said phenomenon is taking place, or at least be able to replicate it. My takeaway is going to the field, trading external validity even at some cost of theoretical precision. Limited to such cases, collecting behavioral evidence to make sense of what is found in the field is more valuable than delving into theoretical discussions on why the framing of tasks might alter results. My research from second chapter onward reflects this shift in perspectives, and my conclusions from the subsequent parts are accordingly more practical.

A key conclusion is the contrast between individual behavior and the institutional settings that constrain it. The positive effects on social capital observed in chapter two emerged from semester-long classroom interactions in small groups with mandatory attendance. This underscores the importance of sustained, equal-status intergroup contact. Yet, such opportunities are the exception, not the rule.[1] Most academic programs offer few electives that foster mixing, meaning students who are capable of making unbiased assessments in diverse settings still find themselves in segregated networks. Ultimately, structural factors that limit meaningful cross-SES connections loom larger than individual preferences. Consequently, policy interventions that target individual behavior alone will likely fail if they do not address these underlying institutional barriers.

Looking forward, this research opens several avenues for future investigation. The challenges encountered in studying digital nudges suggest a need for approaches that capture the interaction between interface design and content personalization. Field experiments that manipulate design features within existing platforms through browser extensions or modified mobile applications, while challenging to implement and problematic because the risk of substitution effects (where users may switch to other platforms), could provide more externally valid and policy-relevant insights.

Regarding the accumulation of social capital, future research should test institutional arrangements designed to maximize cross-group interaction. For example, experimentally increasing the mixed-program courses, allowing for more electives courses across programs, or even strategically allocating classroom spaces for different programs could be cost-effective ways to promote beneficial cross-SES interactions. Information provision interventions, concerning access and eligibility to existing student aid schemes also look promising for future research.

As society grapples with growing economic inequality and our decisions increasingly made

[1] To our knowledge, there exists only two such courses at our partner institution.

in digital environments, I believe this dissertation provides systematic ways to study these realities from an economic perspective. By examining behavior through the lens of structural constraints rather than focusing solely on the individual, my hope is to contribute in designing policies that help regulate markets and institutions toward more equitable outcomes.