

1 Class differences in social networks: Evidence from a referral 1
2 experiment 2

3 Manuel Munoz*, Ernesto Reuben†*, Reha Tuncer‡ 3

4 July 30, 2025 4

5 **Abstract** 5

6 Economic connectivity, defined as the share of high-SES connections in one's network, 6
7 is a strong correlate of labor market income. Yet, low-SES individuals are typically at 7
8 a disadvantage when it comes to knowing the right people. Referral hiring leverages 8
9 networks and make explicit the role of economic connectivity where taste-based biases 9
10 could further exacerbate low-SES outcomes. We conduct a field experiment with 734 10
11 university students to study the network compositions of different SES groups. We 11
12 leverage enrollment networks to identify all potential referral candidates and conduct 12
13 an incentivized referral exercise to reveal SES biases within these choice sets. We find 13
14 that the university enrollment networks are highly segregated, with low-SES and high- 14
15 SES individuals having a higher share of same-SES connections in their networks due 15
16 to program selection (12% and 31% respectively). When considering ex post actualized 16
17 choice sets for the observed referrals, the segregation becomes worse: Low-SES individu- 17
18 als connect with other low-SES individuals at rates 30% higher than the university share, 18
19 while high-SES individuals connect with other high-SES individuals at rates 55% higher 19

*Luxembourg Institute of Socio-Economic Research

†Division of Social Science, New York University Abu Dhabi

‡University of Luxembourg

20 than the university share. Yet, we find no bias against low-SES individuals once we 20
21 account for network structures. We randomly assign half of the participants to a condi- 21
22 tion where their referral candidate receives a fixed bonus on top of pay-for-performance 22
23 referral incentives. We find that additional incentives for the referral candidate do not 23
24 change social proximity with the referral nor the referral quality. Our findings suggest 24
25 that systematic segregation patterns in networks that alter choice sets matter more than 25
26 taste-based SES biases in referrals, and highlight the potential for institutional action 26
27 in promoting SES diversity. 27

28 **JEL Classification:** C93, J71, D85, Z13 28

29 **Keywords:** social capital, social networks, referral hiring, socioeconomic status, field 29
30 experiment 30

31 **1 Introduction**

31

32 Equally qualified individuals in terms of productivity face different labor market out-
33 comes based on their socioeconomic status ([Stansbury & Rodriguez, 2024](#)). This per-
34 sistent inequality undermines meritocratic ideals and represents a substantial barrier to
35 economic mobility. A key driver of SES-based inequality in the labor market stems from
36 differences in social capital.¹ Economic connectivity, defined as the share of high-SES
37 connections among low-SES individuals, is the most important facet of social capital
38 because it correlates strongly with labor market income ([Chetty et al., 2022a](#)). In this
39 sense, a lack of social capital means lack of access to individuals with influential (higher
40 paid) jobs and job opportunities. It implies having worse outcomes when using one's
41 network to find jobs conditional on the capacity to leverage one's social network.²

42 Referral hiring—the formal or informal process where firms ask workers to recom-
43 mend qualified candidates for job opportunities—is a common labor market practice
44 that makes differences in social capital evident.³ Since referrals originate from the net-
45 works of referrers, the composition of referrer networks becomes a crucial channel that
46 propagates inequality. Similar individuals across socio-demographic characteristics form
47 connections at higher rates ([McPherson et al., 2001](#)), making across-SES (low-to-high)
48 connections less likely than same-SES connections ([Chetty et al., 2022a](#)). Referrals will
49 thus reflect similarities in socio-demographic characteristics present in networks even in
50 the absence of biases in the referral procedure—that is, even when referring randomly
51 from one's network according to some productivity criteria.

52 Yet, experimental evidence shows referrals can be biased even under substantial

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

²See 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.

³Referrals solve some frictions in the search and matching process and benefit both job-seekers and employers. As a consequence, referral candidates get hired more often, have lower turnover, and earn higher wages ([Brown et al., 2016](#); [Dustmann et al., 2016](#); [Friebel et al., 2023](#)).

53 pay-for-performance incentives beyond what is attributable to differences in network 53
54 compositions, at least in the case of gender (Beaman et al., 2018; Hederos et al., 2025). 54
55 A similar bias against low-SES individuals may further exacerbate their outcomes. If 55
56 job information is in the hands of a select few high-SES individuals to whom low-SES 56
57 individuals already have limited network access due to their lack of economic connec- 57
58 tivity, and high-SES referrers are biased against low-SES individuals—referring other 58
59 high-SES individuals at higher rates than their network composition would suggest—we 59
60 should expect referral hiring to further disadvantage low-SES individuals. 60

61 The empirical question we answer in this paper is whether referrers exhibit bias 61
62 against low-SES peers after accounting for differences in network SES composition. We 62
63 also evaluate the causal impact of two different incentive structures on referral behavior. 63

64 In this study, we examine inequalities related to SES by curating a university-wide 64
65 network dataset comprising over 4,500 students for whom classroom interactions are 65
66 recorded along with individual attributes. We focus on the role of SES in referrals 66
67 by experimentally investigating whether individuals who are asked to refer a peer tend 67
68 to refer a same-SES candidate. We also explore potential mechanisms behind referral 68
69 patterns by randomizing participants into two different incentive structures. To this end, 69
70 we conducted a lab-in-the-field experiment with 734 students at a Colombian university. 70
71 We instructed participants to refer a qualified student for tasks similar to the math and 71
72 reading parts of the national university entry exam (equivalent to the SAT in the US 72
73 system). To incentivize participants to refer qualified candidates during the experiment, 73
74 we set earnings to depend on referred candidates' actual university entry exam scores. 74

75 Referral hiring in the labor market can range from firm-level formal referral programs 75
76 asking employees to bring candidates to simply passing on job opportunities between 76
77 network members (Topa, 2019). Since our participants are students at the university 77
78 and refer based on exam scores, we abstract away from formal referral programs with 78
79 defined job openings. Our setting instead resembles situations where contacts share 79
80 opportunities with each other without requiring the referred candidate to take any action 80
81 and without revealing the referrer's identity. This eliminates reputational concerns since 81

82 there is no hiring employer. It also establishes a lower bound on the expected reciprocity 82
83 for the referrer when combined with pay-for-performance incentives (Bandiera et al., 83
84 2009; Witte, 2021). At the same time, referring based on university entry exam scores 84
85 is still an objective, widely accepted measure of ability. We show evidence that referrers 85
86 in our setting not only possess accurate information about these signals but can also 86
87 screen more productive individuals from their university network. 87

88 In a university setting, class attendance provides essential opportunities for face- 88
89 to-face interaction between students. This is a powerful force that reduces network 89
90 segregation by providing ample opportunities to meet across SES groups, because of 90
91 exposure to an equal or higher level of high-SES individuals compared to the general 91
92 population (Chetty et al., 2022b).⁴ The very high level of income inequality in Colombia 92
93 makes SES differences extremely visible in access to tertiary education, where rich and 93
94 poor typically select into different institutions (Jaramillo-Echeverri & Álvarez, 2023). 94
95 However, in the particular institutional setting we have chosen for this study, different 95
96 SES groups mix at this university, allowing us to focus on SES diversity within the 96
97 institution. At the same time, as students take more classes together, their similarities 97
98 across all observable characteristics tend to increase (Kossinets & Watts, 2009). This 98
99 is an opposite force that drives high- and low-SES networks to segregate. We observe 99
100 the net effect of these two opposing forces using administrative data and construct class 100
101 attendance (enrollment) networks for 734 participants based on the number of common 101
102 courses they have taken together with other students. This allows us to directly identify 102
103 aggregate characterizations of different SES groups' network compositions as a function 103
104 of courses taken (e.g., in same-SES share), as well as the individual characteristics of 104
105 network members who receive referrals among all possible candidates. 105

106 We find strong evidence that networks of high- and low-SES participants exhibit 106
107 same-SES bias. On average, both groups connect with their own SES group at higher 107

⁴In a different sample from the same university population, Díaz et al. (2025) show this holds true for the highest-SES individuals at this institution, accounting for about 6% of their sample but less than 5% of Colombian high-school graduates (Fergusson & Flórez, 2021a).

108 rates than would occur randomly given actual group shares at the university (12% for 108
109 low-SES and 31% for high-SES). As students take more courses together within the 109
110 same program, their networks dwindle in size and become even more homogeneous in 110
111 SES shares. At 12 courses together (the median number of courses taken together among 111
112 referrals), the same-SES share increases to 30% above the university share for low-SES 112
113 students and 55% above for high-SES students. We identify selection into academic 113
114 programs as a key mechanism explaining this phenomenon: The private university where 114
115 our study took place implements exogenous cost-based program pricing and does not offer 115
116 SES-based price reductions. This results in programs with very large cost differences 116
117 within the same university, with some programs costing up to six times the cheapest 117
118 one. We find that the average yearly fee paid per student increases with SES, and the 118
119 high-SES share in the most expensive program at the university—medicine—drives a 119
120 large part of the network segregation across SES groups. 120

121 Do segregated networks account for the differences in SES referral rates across SES 121
122 groups? Same-SES referrals are 17% more common than referrer networks suggest. 122
123 Controlling for differences in network compositions, we find that the entirety of the bias 123
124 is driven by low-SES referrers. We find no bias against low-SES peers beyond what is 124
125 attributable to differences in network composition. Regardless of SES, participants refer 125
126 productive individuals, and referred candidates are characterized by a very high number 126
127 of courses taken together. The latter underlies the impact of program selection on the 127
128 intensity of social interaction, where participants activate smaller and more homogeneous 128
129 parts of their networks for making referrals. Our treatment randomized participants 129
130 across two different incentive schemes by adding a substantial monetary bonus (\$25) 130
131 for the referred candidate on top of the pay-for-performance incentives. We provide 131
132 evidence that treatment incentives did not change referral behavior across the same-SES 132
133 referral rate, the number of courses taken together with the referral candidate, and the 133
134 candidate's exam scores. We interpret the lack of differences in the number of courses 134
135 taken together as further evidence that referrals go to strong social ties across both 135

¹³⁶ treatments regardless of the incentive structure.⁵ ¹³⁶

¹³⁷ Our main empirical contribution to the experimental referral literature is our obser- ¹³⁷
¹³⁸ vation of the entire network that characterizes the referral choice set. Earlier research ¹³⁸
¹³⁹ compares referrals made across different incentive structures and makes inferences about ¹³⁹
¹⁴⁰ the counterfactual. For example, [Beaman and Magruder \(2012\)](#) compared referrers paid ¹⁴⁰
¹⁴¹ based on their referred candidate's productivity instead of receiving a fixed finder's fee, ¹⁴¹
¹⁴² and [Beaman et al. \(2018\)](#) compared referrers who were restricted to refer either a male ¹⁴²
¹⁴³ or female candidate instead of choosing freely. While [Pallais and Sands \(2016\)](#) recruited ¹⁴³
¹⁴⁴ a random sample of non-referred workers for comparison with referred ones, none of ¹⁴⁴
¹⁴⁵ the previous studies could identify the entire referral choice set and provide a direct ¹⁴⁵
¹⁴⁶ comparison to those who were referred by the participants. Observing the entire net- ¹⁴⁶
¹⁴⁷ work allows us to identify biases in referrals in a more natural way, without imposing ¹⁴⁷
¹⁴⁸ restrictions on the choice sets. A similar approach to ours is [Hederos et al. \(2025\)](#), who ¹⁴⁸
¹⁴⁹ elicited friendship networks by asking referrers to name 5 close friends. Their findings ¹⁴⁹
¹⁵⁰ suggest only half of those who were referred were from the elicited friendship network, ¹⁵⁰
¹⁵¹ and thus represent an incomplete observation of the entire referral choice set. We take ¹⁵¹
¹⁵² our analysis one step further by requesting referrals from the enrollment network, where ¹⁵²
¹⁵³ we have complete information on every single connection that may or may not receive ¹⁵³
¹⁵⁴ a referral. This allows us to neatly separate the effect of network composition from any ¹⁵⁴
¹⁵⁵ potential biases stemming from the referral procedure itself. ¹⁵⁵

¹⁵⁶ Second, we build upon the earlier work on inequalities in referrals and the role of SES ¹⁵⁶
¹⁵⁷ differences. The reliance of labor markets on referrals, coupled with homophily in social ¹⁵⁷
¹⁵⁸ networks, can lead to persistent inequalities in wages and employment ([Bolte et al., 2021](#); ¹⁵⁸
¹⁵⁹ [Calvo-Armengol & Jackson, 2004](#); [Montgomery, 1991](#)). The premise of these models is ¹⁵⁹
¹⁶⁰ that referrals exhibit homophily, so that employees are more likely to refer workers of ¹⁶⁰
¹⁶¹ their own race, gender, SES, etc. Supporting evidence shows that low-SES individuals ¹⁶¹
¹⁶² have networks with lower shares of high-SES individuals, which partly explains why they ¹⁶²

⁵This follows directly from earlier evidence showing that referrals tend to go to strong ties, i.e., close friends and/or family members ([Gee et al., 2017](#); [Kramarz & Nordström Skans, 2014](#); [Wang, 2013](#)).

163 have worse labor market outcomes (Chetty et al., 2022a; Stansbury & Rodriguez, 2024). 163
164 We contribute by separately identifying the role of network homophily (the tendency 164
165 to connect with similar others) and referral homophily (the tendency to refer similar 165
166 others). Our results suggest that network homophily, rather than referral homophily, 166
167 drives SES inequality in our setting. 167

168 To our knowledge, Díaz et al. (2025) are the first to study SES biases in referrals, 168
169 and our study is conceptually the closest to theirs. Drawing from a similar sample at 169
170 the same institution, Díaz et al. (2025) focus on referrals from first-year students made 170
171 within mixed-program classrooms and find no evidence for an aggregate bias against low- 171
172 SES individuals. We also find no aggregate bias against low-SES individuals in referrals 172
173 beyond what is attributable to differences in network structure. Our setup differs as we 173
174 sample from students who completed their first year and impose no limits on referring 174
175 from a classroom. This has several implications: We find that referrals in our setup go to 175
176 individuals within the same program, and that programs have different SES shares which 176
177 become even more accentuated as students take more courses together. While networks 177
178 drive inequality in referral outcomes because of the institutional environment in our 178
179 sample, we have no reason to believe first-year student networks in Díaz et al. (2025) 179
180 have similar levels of segregation to begin with. Our findings suggest that implementing 180
181 more mixed-program courses that allow for across-SES mixing should be a clear policy 181
182 goal to reduce segregation (Alan et al., 2023; Rohrer et al., 2021). 182

183 The remainder of the paper is organized as follows. Section 2 begins with the back- 183
184 ground and setting in Colombia. In Section 3 we present the design of the experiment. 184
185 In Section 4 we describe the data and procedures. Section 6 discusses the results of 185
186 the experiment and Section 7 introduces robustness checks. Section 8 concludes. The 186
187 Appendix presents additional tables and figures as well as the experiment instructions. 187

¹⁸⁸ 2 Background and Setting

¹⁸⁸

¹⁸⁹ Our experiment took place in Colombia, a country that consistently ranks highly in ¹⁸⁹
¹⁹⁰ terms of economic inequality. The richest decile of Colombians earn 50 times more than ¹⁹⁰
¹⁹¹ the poorest decile ([United Nations, 2023](#); [World Bank, 2024](#)). This economic disparity ¹⁹¹
¹⁹² creates profound differences in outcomes across SES groups in terms of education, geo- ¹⁹²
¹⁹³ graphic residence, language, manners, and social networks ([Angulo et al., 2012](#); [García ¹⁹³](#)
¹⁹⁴ et al., 2015; [García Villegas & Cobo, 2021](#)). While these patterns are not atypical and ¹⁹⁴
¹⁹⁵ exist elsewhere, Colombia's pronounced inequality makes economic, educational, and ¹⁹⁵
¹⁹⁶ cultural differences across SES particularly visible and thus provides an ideal setting to ¹⁹⁶
¹⁹⁷ study SES biases in referral selection. ¹⁹⁷

¹⁹⁸ We rely on Colombia's established estrato classification system to measure SES in ¹⁹⁸
¹⁹⁹ our study. In 1994, Colombia introduced a nationwide system that divides the popula- ¹⁹⁹
²⁰⁰ tion into six strata based on "similar social and economic characteristics" ([Hudson & ²⁰⁰](#)
²⁰¹ Library of Congress, 2010, p. 102). Designed for utility subsidies from higher strata to ²⁰¹
²⁰² support lower strata, the system aligns with and reinforces existing social class divisions ²⁰²
²⁰³ ([Guevara S & Shields, 2019](#); [Uribe-Mallarino, 2008](#)). It is widely used by policymakers ²⁰³
²⁰⁴ and in official statistics ([Fergusson & Flórez, 2021a](#)). Using the estrato system, we cat- ²⁰⁴
²⁰⁵ egorize students in strata 1-2 as low-SES, strata 3-4 as middle-SES, and strata 5-6 as ²⁰⁵
²⁰⁶ high-SES. ²⁰⁶

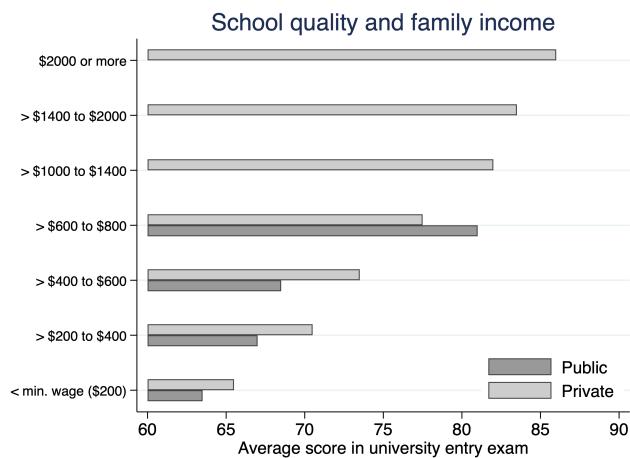
²⁰⁷ Colombia's educational segregation typically prevents meaningful interaction be- ²⁰⁷
²⁰⁸ tween socioeconomic groups, as wealthy families attend exclusive private schools while ²⁰⁸
²⁰⁹ poorer families access lower-quality public or "non-elite" private institutions (see Figure ²⁰⁹
²¹⁰ 1). Our study takes place in a non-elite private university which attracts students across ²¹⁰
²¹¹ the socioeconomic spectrum: The university's student body comprises 35% low-SES, ²¹¹
²¹² 50% middle-SES, and 15% high-SES students.⁶ This diversity provides opportunities ²¹²
²¹³ for different SES groups to meet and interact within the same institutional framework. ²¹³

²¹⁴ The partner university creates conditions for contact on equal status. All students ²¹⁴

⁶Government statistics reveal less than 5% of the population is high-SES ([Hudson & Library of Congress, 2010](#), p. 103).

215 pay the same fees based on their program choices, and less than 5% of students receive 215
216 scholarships. The student body is mostly urban and has comparable university entry 216
217 exam scores due to the entrance exam **ADD NUMBERS**. These additional factors make 217
218 our setting appropriate to study the effects of contact on intergroup discrimination. 218

Figure 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 score in the 65-70 band, the private university where we conducted this experiment caters to low- and high-SES students. Figure reproduced from [Fergusson and Flórez \(2021b\)](#).

219 We construct enrollment networks using administrative data to map social connec- 219
220 tions at the university. For each participant, we identify all other undergraduate students 220
221 with whom they have taken at least one course and create their individual network of 221
222 university connections. The size of this network depends on how many different students 222
223 a participant has encountered through coursework, while the intensity of connection is 223
224 measured by the number of courses taken together. This approach provides a complete 224
225 picture of each participant's social environment at the university, including detailed 225
226 characteristics (i.e., SES, academic program, performance) for both the participant and 226
227 every person in their network. 227

228 why don't we introduce number of programs and student count here as background? 228

229 **3 Empirical Specification**

229

230 To formally test for SES bias, we use a choice modeling approach. We observe a single re- 230
 231 ferral outcome from mutually exclusive candidates. Our design leverages the enrollment 231
 232 network to generate a dataset which includes alternative-specific variables for each refer- 232
 233 ral decision, i.e., SES, courses taken together with the participant making the referral, as 233
 234 well as entry exam scores for not just the chosen alternative but all referral candidates. 234
 235 Using a conditional logit model, we can identify whether an SES group has an aggregate 235
 236 bias controlling for each individual's unique enrollment network composition. 236

237 We follow an additive random utility model framework where individual i and alter- 237
 238 native j have utility U_{ij} that is the sum of a deterministic component, V_{ij} , that depends 238
 239 on regressors and unknown parameters, and an unobserved random component ε_{ij} : 239

240 We observe the outcome $y_i = j$ if alternative j has the highest utility of the alterna- 240
 241 tives. The probability that the outcome for individual i is alternative j , conditional on 241
 242 the regressors, is: 242

$$p_{ij} = \Pr(y_i = j) = \Pr(U_{ij} \geq U_{ik}), \quad \text{for all } k \quad (1)$$

243 The conditional logit model specifies that the probability of individual i choosing 243
 244 alternative j from choice set C_i is given by: 244

$$p_{ij} = \frac{\exp(x'_{ij}\beta)}{\sum_{l \in C_i} \exp(x'_{il}\beta)}, \quad j \in C_i \quad (2)$$

245 where x_{ij} are alternative-specific regressors, i.e., characteristics of potential referral 245
 246 candidates that vary across alternatives. In our context, individual i chooses to refer 246
 247 candidate j from their enrollment network C_i . The alternative-specific regressors include 247
 248 SES and entry exam scores of the referral candidate, and the number of courses taken 248
 249 together with the participant making the referral. Conditional logit structure eliminates 249
 250 participant-specific factors that might influence both network formation and referral 250
 251 decisions, allowing us to identify preferences within each participant's realized network. 251

252 For causal identification of SES bias, we require two identifying assumptions. Specif- 252
253 ically: 253

254 1. **Conditional exogeneity.** SES and the number of courses taken together could 254
255 be endogenous due to program selection. High-SES students sort into expensive 255
256 programs while low-SES students choose affordable programs, creating systematic 256
257 SES variation across enrollment networks. Similarly, the number of courses taken 257
258 together reflects program selection decisions that may correlate with unobserved 258
259 referral preferences. However, conditional on the realized enrollment network, the 259
260 remaining variation in both SES and the number of courses taken together across 260
261 referral candidates must be independent of unobserved factors affecting referral 261
262 decisions. In the robustness checks, we show that being in the same program 262
263 with the referrer does not impact our SES bias estimates, although it reduces the 263
264 coefficient on the number of courses taken together. 264

265 2. **Complete choice sets and independence of irrelevant alternatives.** Ad- 265
266 ministrative data captures the complete enrollment network, with all peers who 266
267 took at least one course with individual i and represent the true choice set for re- 267
268 ferral decisions (unless participants have potential referral candidates with whom 268
269 they never took classes). The independence of irrelevant alternatives (IIA) as- 269
270 sumption requires that choices between any two alternatives be independent of 270
271 other options in the choice set, which could be problematic if, e.g., peers within 271
272 the same SES group are viewed as close substitutes. This concern does not apply 272
273 to our setting because the design of our experiment ensures that choice sets are 273
274 fixed by enrollment rather than arbitrary inclusion/exclusion of alternatives that 274
275 create IIA violations. 275

276 Under these assumptions, the conditional logit framework controls for individual het- 276
277 erogeneity in program selection (absorbed by conditioning on choice sets), selection into 277
278 programs based on observable characteristics (through alternative-specific variables), and 278
279 choice set composition effects (through the multinomial structure). Therefore, β should 279

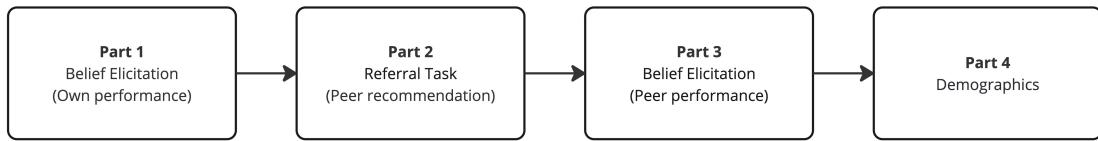
280 identify the causal effect of referral candidate SES on referral probability, holding con- 280
281 stant the number of courses taken together and the entry exam scores of candidates. A 281
282 significant coefficient will then indicate taste-based discrimination. 282

283 We pool participants by their SES group, and estimate the above described con- 283
284 ditional fixed effects logit model once for low-, middle-, and high-SES referrers. We 284
285 standardize entry exam scores and the number of courses taken together at the individ- 285
286 ual network level. For each referrer’s network, we first calculate the mean and standard 286
287 deviation for both measures. We then compute the average of these means and standard 287
288 deviations across all 734 referrers. Each referral candidate’s entry exam score and the 288
289 number of courses they have taken with the referrer is standardized using these network- 289
290 level statistics. The standardization formula is $z_i = (x - \bar{X}_i)/\sigma_i$, where \bar{X}_i and σ_i are 290
291 the average of network means and standard deviations for C_i . 291

292 4 Design 292

293 We designed an online experiment to assess peer referral selection from an SES perspec- 293
294 tive and to evaluate the causal effect of providing a bonus to referral candidates. The 294
295 experimental design consisted of two incentivized tasks administered in the following 295
296 sequence: First, participants completed belief elicitation tasks about their own perfor- 296
297 mance on the national university entry exam. Second, they completed the main referral 297
298 task, nominating peers based on exam performance in two academic areas. Finally, par- 298
299 ticipants reported beliefs about their referrals’ performance and provided demographic 299
300 information. This structure allowed us to measure both the accuracy of participants’ 300
301 beliefs and their referral behavior under controlled incentive conditions. Figure 2 shows 301
302 the experimental timeline, and detailed instructions are provided in Appendix B. 302

Figure 2: Experimental Timeline



Note: Participants first reported beliefs about their own university entry exam performance, then made referrals for each academic area, and finally reported beliefs about their referrals' performance and provided demographics.

303 4.1 Performance measures

304 To establish an objective basis for referral performance, we use national university entry 304
 305 exam scores (SABER 11). All Colombian high school students take the SABER 11 305
 306 exam at the end of their final year as a requirement for university admission. The scores 306
 307 from this exam provide pre-existing, comparable measures of performance. By using 307
 308 existing administrative data, we also ensure that all eligible students have comparable 308
 309 performance measures. 309

310 The exam consists of five areas (critical reading, mathematics, natural sciences, social 310
 311 sciences, and English). We focus on critical reading and mathematics as these repre- 311
 312 sent two independent and fundamental skills. Critical reading evaluates competencies 312
 313 necessary to understand, interpret, and evaluate texts found in everyday life and broad 313
 314 academic fields (e.g., history). Mathematics assesses students' competency in using un- 314
 315 dergraduate level mathematical tools (e.g., reasoning in proportions, financial literacy). 315
 316 These together capture performance in comprehending and critically evaluating written 316
 317 material as well as reasoning and problem-solving abilities. 317

318 For each area, we calculate percentile rankings based on the distribution of scores 318
 319 among all currently enrolled students, providing a standardized measure of relative per- 319
 320 formance within the university population. 320

321 **4.2 Referral task**

321

322 The main task involves making referrals among peers. For both exam areas (critical 322
323 reading and mathematics), participants refer one peer they believe excels in that area. 323
324 We provide an example question from the relevant exam area to clarify the skills that 324
325 are being assessed. Participants type the name of their preferred candidate to make 325
326 a referral. To avoid issues with recall, the interface provides autocomplete name and 326
327 program suggestions from the administrative database (see Figure 3). 327

Figure 3: Referral task interface

Your recommendation

We are interested in your recommendation of the person you consider
best to solve similar problems to those in the **Math test**.

- * Only someone with whom you have taken at least one class...
- * We will not contact your recommendation...

Please write the name of your recommendation:

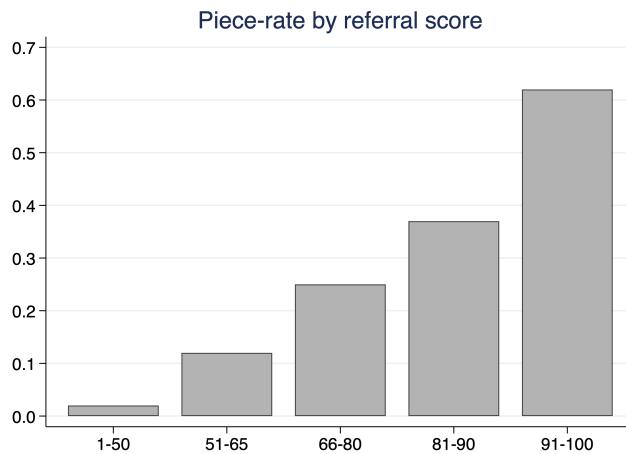
John
John Lennon (Music - 2018) 
John Stuart Mill (Law - 2020)

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

328 Participants can only refer students with whom they have taken at least one class 328
329 during their university studies. This requirement ensures that referrals are based on 329
330 actual peer interactions. We randomize the order in which participants make referrals 330
331 across the two exam areas. 331

332 We incentivize referrals using a piece rate payment structure. Referrers earn in- 332
333 creasing payments as the percentile ranking of their referral increases (see Figure 4). We 333
334 multiply the piece rate coefficient associated with the percentile rank by the actual exam 334
335 scores of the referral to calculate earnings. This payment structure provides strong in- 335
336 centives to refer highly ranked peers with potential earnings going up to \$60 per referral. 336

Figure 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.

338 4.3 Bonus Treatment 338

339 To examine how different incentive structures affect referral selection, we randomly assign 339
 340 a fixed bonus payment for students who get a referral. In the **Baseline** treatment, only 340
 341 the participants can earn money based on their referral's performance. The **Bonus** 341
 342 treatment adds an additional fixed payment of \$25 to the peer who gets the referral. 342
 343 This payment is independent of the referral's actual performance (see Table 1). 343

Table 1: Incentive structure by treatment

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

⁷Note 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.

344 We use a between-subjects design and randomly assign half our participants to the 344
345 **Bonus** treatment. This allows us to causally identify the effect of the bonus on referral 345
346 selection. Participants learn whether their referral gets the fixed bonus before making 346
347 referral decisions. 347

348 4.4 Belief elicitation 348

349 We collect two sets of incentivized beliefs to assess the accuracy of participants' knowl- 349
350 edge about exam performance. Participants first report beliefs about their own percentile 350
351 ranking in the university for each exam area. After making referrals, participants report 351
352 their beliefs about their referrals' percentile ranking in the university. For both belief 352
353 elicitation tasks, participants earn \$5 per correct belief if their guess is within 7 per- 353
354 centiles of the true value. This margin of error is designed to balance precision with the 354
355 difficulty of the task. 355

356 5 Sample, Incentives, and Procedure 356

357 We invited all 4,417 undergraduate students who had completed their first year at the 357
358 university at the time of recruitment to participate in our experiment. A total of 837 358
359 students participated in the data collection (19% response rate). Our final sample con- 359
360 sists of 734 individuals who referred peers with whom they had taken at least one class 360
361 together (88% success rate). We randomly allocated participants to either **Baseline** or 361
362 **Bonus** treatments. 362

363 Table 2 presents key demographic characteristics and academic performance indi- 363
364 cators across treatments (see Appendix Table A.1 for selection). The sample is well- 364
365 balanced between the **Baseline** and **Bonus** conditions and we observe no statistically 365
366 significant differences in any of the reported variables (all p values > 0.1). Our sample is 366
367 characterized by a majority of middle-SES students with about one-tenth of the sample 367
368 being high-SES students. The test scores and GPA distributions are balanced. On av- 368
369 erage, participants had taken 3.8 courses together with members of their network, and 369

370 the average network consisted of 175 peers.

370

Table 2: Balance between treatments

	Baseline	Bonus	<i>p</i>
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 refers to the average number of network members. Low-SES, Med-SES, and High-SES indicate SES categories based on strata.

371 The experiment was conducted online through Qualtrics, with participants recruited 371
372 from active students. To ensure data quality while managing costs, we randomly selected 372
373 one in ten participants for payment. Selected participants received a fixed payment of 373
374 \$17 for completion. They also received potential earnings from one randomly selected 374
375 belief question (up to \$5) and one randomly selected referral question (up to \$60). This 375
376 structure resulted in maximum total earnings of \$82. The average time to complete 376
377 the survey was 30 minutes, with an average compensation of \$80 for the one in ten 377
378 participants randomly selected for payment. Payment processing occurred through bank 378
379 transfer within 15 business days of participation. 379

380 **6 Results**

380

381 **6.1 Network characteristics**

381

382 We begin by describing the key features of the enrollment networks. On average, par-
383 ticipants connect with 175 other students, and take an average of 3.62 courses together.
384 Figure 5 shows how network characteristics vary by students' time at the university:
385 both the number of connections (network size) and the number of courses taken to-
386 gether (connection intensity) change as participants progress through their studies.

382

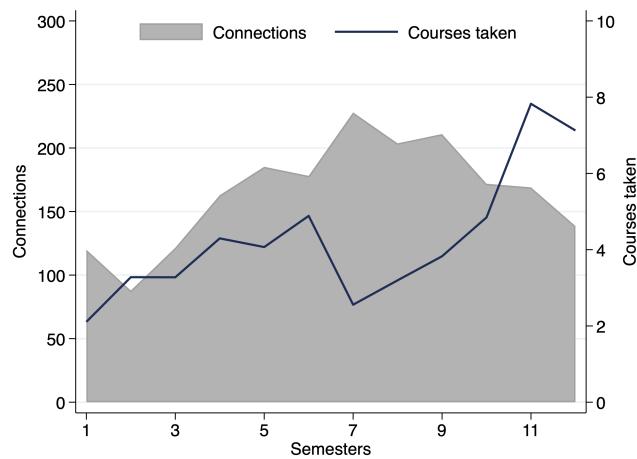
383

384

385

386

Figure 5: Network size and courses taken together by time spent at
the university

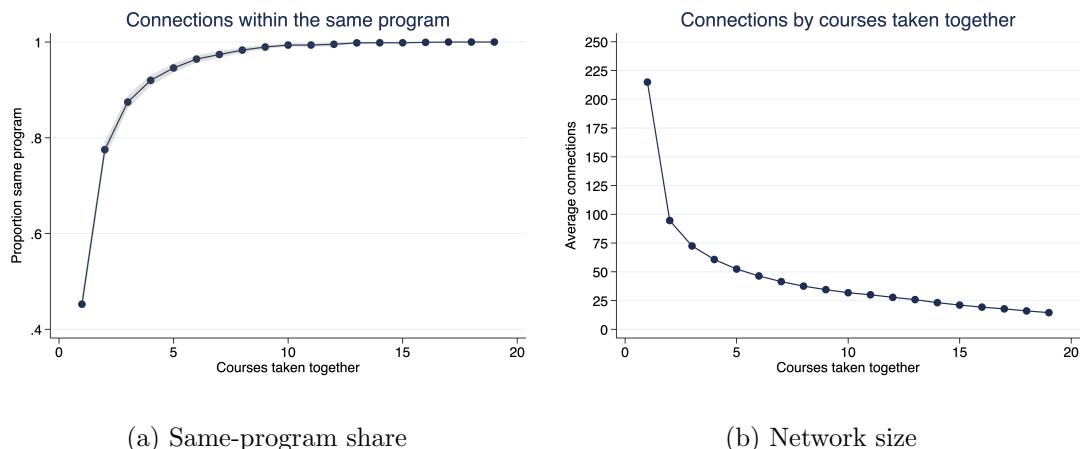


Note: This figure displays the average number of connections in blue and the average number of courses taken together with connections in grey across semesters completed. Network size (nb. of connections) peaks around 7 semesters before declining as students graduate. Connection intensity (nb. of courses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

387 We now examine how connection intensity relates to network size and composition. 387
388 First, if two students take more courses together, it is very likely that they are in the 388
389 same academic program. We plot this relationship in Figure 6a: As students take more 389
390 than 5 courses together, the share of students in their enrollment network from the same 390
391 academic program quickly exceeds 90%. Second, because students sort into specialized 391

392 academic programs, increases in courses taken together should result in decreases in 392
 393 connections. We plot this relationship in Figure 6b: As students take more than 5 393
 394 courses together, the size of their enrollment network drops dramatically from above 210 394
 395 to below 50. These patterns reveal that while participants' overall networks are large 395
 396 with relatively few courses taken together on average, they spend most of their time at 396
 397 the university within smaller, more intensive groups of peers from the same academic 397
 398 program. 398

Figure 6: Network characteristics and courses taken together



(a) Same-program share

(b) Network size

Note: The left panel illustrates the share of connections within the same program as a function of the number of courses taken together. The right panel 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.

399 6.2 Referral characteristics 399

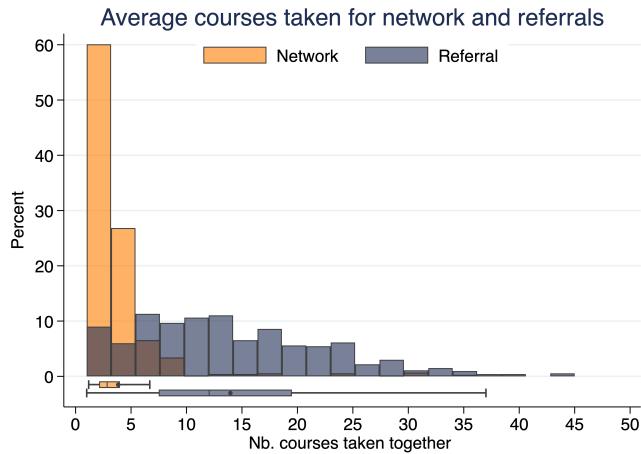
400 Participants made one referral for math and one referral for the reading part of the 400
 401 university entry exam from their enrollment networks. We observe 1,342 referrals from 401
 402 734 participants in our final dataset. More than 90% of these consist of participants 402

403 referring for both exam areas (see Appendix Table A.2). About 70% of these referrals 403
404 go to two separate individuals. We compare the outcomes across exam areas for referrals 404
405 only going to separate individuals in Appendix Table A.3 and all referrals in Appendix 405
406 Table A.4. In both cases, we find no meaningful differences between referrals made for 406
407 Math or Reading areas of the entry exam. As referrals in both exam areas come from 407
408 the same enrollment network, we pool referrals per participant and report their averages 408
409 in our main analysis to avoid inflating statistical power in our comparisons. 409

410 What are the characteristics of the individuals who receive referrals, and how do 410
411 they compare to others in the enrollment network? Because we have an entire pool of 411
412 potential candidates with one referral chosen from it, we compare the distributions for 412
413 our variables of interest between the referred and non-referred students. 413

414 First, referrals go to peers with whom the referrer has taken around 14 courses with 414
415 on average, compared to almost 4 on average with others in their network (see Figure 415
416 7). This difference of 10.1 courses is significant ($t = 34.98, p < 0.001$), indicating 416
417 that referrers choose individuals with whom they have stronger ties. While the median 417
418 referral recipient has taken 12 courses together with the referrer, the median network 418
419 member has shared only 2.8 courses. The interquartile range for referrals spans from 7.5 419
420 to 19.5 courses, compared to just 2.1 to 4.0 courses for the broader network, highlighting 420
421 the concentration of referrals among peers with high social proximity. In addition, 93% 421
422 of referrals go to students within same program. 422

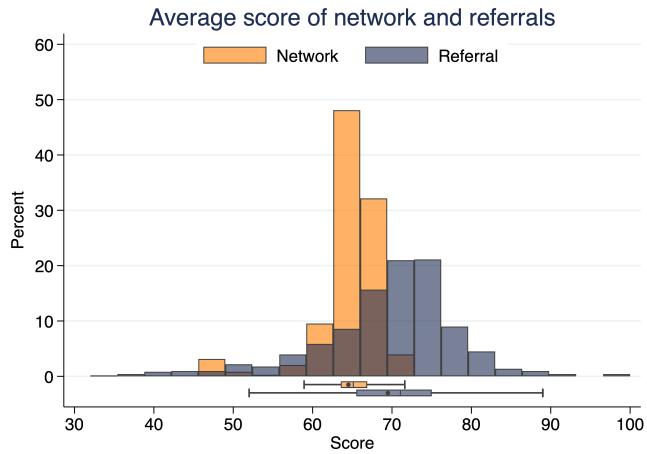
Figure 7: Courses taken together with network members and referrals



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 $D = 33.37$, $p < 0.001$).

423 Second, we examine entry exam score differences between referred students and the 423
 424 broader network. Referrals go to peers with an average score of 69.5 points, compared to 424
 425 64.5 points for other network members (see Figure 8). This difference of 5 points is sig- 425
 426 nificant ($t = 18.97$, $p < 0.001$), indicating that referrers choose higher-performing peers. 426
 427 While the median referral recipient scores 71 points, the median network member scores 427
 428 65.1 points. The interquartile range for referrals spans from 65.5 to 75 points, compared 428
 429 to 63.5 to 66.9 points for the broader network, highlighting the clear concentration of 429
 430 referrals among higher performing peers. 430

Figure 8: Entry exam scores of network members and referrals



Note: This figure compares the distributions of entry exam scores (Math and 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 $D = 71.16$, $p < 0.001$).

431 6.3 Effect of the Bonus treatment

432 Do referred individuals have different outcomes across treatments? We compare the 432 performance, number of courses taken together, and SES shares of referred individuals 433
 434 between the **Baseline** and **Bonus** treatments in Table 3. While performance of referrals 434 across Reading, Math, and GPA are similar across treatments, middle- and high-SES 435 shares have significant differences. We find that referrals under the **Bonus** condition 436 referred a higher proportion of high-SES individuals (13.5% vs 8.8%, $p = 0.041$) and 437 a lower proportion of middle-SES individuals on average (47.0% vs 53.7%, $p = 0.072$). 438 The similarities in academic performance and number of courses taken together suggest 439 that performance and contact intensity drive referrals regardless of treatment. For this 440 reason, in the remainder of the paper, we report pooled results combining the averages 441 of referral outcomes across treatments. 442

Table 3: Characteristics of referrals by treatment condition

	Baseline Referred	Bonus Referred	<i>p</i>
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
Low-SES	0.376	0.395	0.593
Middle-SES	0.537	0.470	0.072
High-SES	0.088	0.135	0.041
Observations	382	352	

Note: This table compares the characteristics of network members who were referred under baseline vs bonus treatment 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 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. Low-SES, Med-SES, and High-SES are binary variables indicating the share of participants in estrato 1-2, 3-4, or 5-6, respectively. Both columns include only network members who were actually nominated for referral in each treatment condition.

6.4 Identifying the SES bias in referrals

We now present our empirical findings and describe our first set of findings in Table 4. To begin with, the variance explained by all three models are extremely low, suggesting the role of potential SES biases in referrals that go beyond the network structure must be limited. Regardless, controlling for network composition, 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: SES bias in referral decisions by referrer SES group

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.453*** (0.109)	-0.019 (0.098)	-0.710** (0.333)
High-SES referral	-0.584*** (0.211)	-0.255* (0.145)	0.001 (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.

452 Next, we include social proximity controls in our analysis. We proceed by adding 452
 453 the standardized number of courses taken together as a control in our specification and 453
 454 describe the results in Table 5. A one standard deviation increase in the number of 454
 455 courses taken together proves to be highly significant across all models, with coefficients 455
 456 ranging from 0.856 to 1.049, indicating that intensity of contact substantially increase 456
 457 the probability of referral. The high χ^2 statistics suggest that these models explain 457
 458 considerably more variance than specifications without this control, highlighting the 458
 459 predictive power of courses taken together in referral decisions. Nevertheless, low-SES 459
 460 participants still show a strong same-SES bias relative to referring middle-SES peers 460
 461 at the average number of courses taken together. This same-SES bias is not observed 461
 462 among middle-SES or high-SES referrers, who also display no statistically significant 462

⁴⁶³ bias toward low-SES candidates. No referrer group shows a positive bias for high-SES ⁴⁶³
⁴⁶⁴ candidates relative to middle-SES candidates. ⁴⁶⁴

Table 5: SES bias in referral decisions by referrer SES group

	Referrer SES		
	Low	Middle	High
	(1)	(2)	(3)
Low-SES referral	0.348*** (0.123)	-0.064 (0.115)	-0.489 (0.337)
High-SES referral	-0.366 (0.223)	-0.165 (0.157)	-0.140 (0.286)
Courses taken (z-score)	0.856*** (0.035)	0.931*** (0.037)	1.049*** (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.

⁴⁶⁵ We add standardized entry exam scores (Math and Reading average) as a second ⁴⁶⁵
⁴⁶⁶ control variable and describe our results in Table 6. A one standard deviation increase ⁴⁶⁶
⁴⁶⁷ in the entry exam score proves highly significant across all models, with coefficients ⁴⁶⁷
⁴⁶⁸ ranging from 0.587 to 0.883. This shows merit-based considerations due to the incentive ⁴⁶⁸
⁴⁶⁹ structure of the experiment remained central to referral decisions. The slightly higher χ^2 ⁴⁶⁹
⁴⁷⁰ statistics compared to the earlier specification suggests that entry exam scores improve ⁴⁷⁰
⁴⁷¹ model fit. The inclusion of standardized entry exam scores strengthens SES biases. Low- ⁴⁷¹

472 SES referrers maintain their same-SES bias, with now a significant negative bias against 472
473 high-SES. Middle-SES referrers, previously showing no SES bias, now show marginal 473
474 negative bias against high-SES. Finally, high-SES referrers exhibit marginal negative 474
475 bias against low-SES candidates. 475

476 The evidence of a bias becoming significant when controlling for entry exam scores 476
477 has a nuanced interpretation. While at the university-level, low-SES typically score 477
478 lower in the entry exam, low-SES students appearing in high-SES networks are posi- 478
479 tively selected, scoring about 0.14 standard deviations higher than middle-SES students 479
480 (see Appendix Table A.5). Controlling for performance thus removes this positive se- 480
481 lection and reveals the “pure” SES bias that was previously underestimated by above 481
482 average performance of low-SES. Vice versa, high-SES in low-SES networks perform 482
483 0.12 standard deviations better than middle-SES students. The same bias was underes- 483
484 timated as high-SES candidates’ better performance relative to middle-SES in the same 484
485 networks provided a meritocratic justification for getting more referrals. Controlling for 485
486 exam scores reveal that both high- and low-SES referrers have negative SES bias towards 486
487 one another that operates independently of - and counter to - performance-based con- 487
488 siderations. What makes interpretation difficult is that while biased against low-SES, 488
489 high-SES referrers do not (under any specification) display a positive bias towards their 489
490 in-group. For this final reason, we do not dig any further in this direction. 490

491 To conclude, we conduct joint significance tests, testing whether low- and high-SES 491
492 regression coefficients are jointly different from middle-SES for each regression specifi- 492
493 cation. For low-SES referrers, the joint test remains highly significant across all three 493
494 specifications ($\chi^2 = 10.20, p = 0.006$ in the final model), indicating persistent SES bias 494
495 across all specifications. In contrast, middle-SES referrers display no significant joint 495
496 SES bias in any specification, with the test becoming increasingly non-significant as 496
497 controls are added ($\chi^2 = 4.13, p = 0.127$ in the final model). High-SES referrers simi- 497
498 larly show no significant joint SES bias across all three models ($\chi^2 = 4.28, p = 0.118$ in 498
499 the final model). These results suggest that SES bias in referrals is primarily driven by 499
500 low-SES. There is no sufficient evidence to conclude that middle- and high-SES referrers 500

501 systematically discriminate against other-SES peers. Naturally, this null result occurs 501
 502 once we take into account the potential differences in the network compositions of each 502
 503 SES group. In the next section, we explore the differences in the network compositions 503
 504 of different SES groups. 504

Table 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** (0.123)	-0.159 (0.114)	-0.600* (0.327)
High-SES referral	-0.445** (0.222)	-0.274* (0.157)	-0.345 (0.287)
Courses taken (z-score)	0.859*** (0.036)	0.948*** (0.038)	1.043*** (0.118)
Entry exam (referral z-score)	0.607*** (0.052)	0.587*** (0.047)	0.883*** (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.

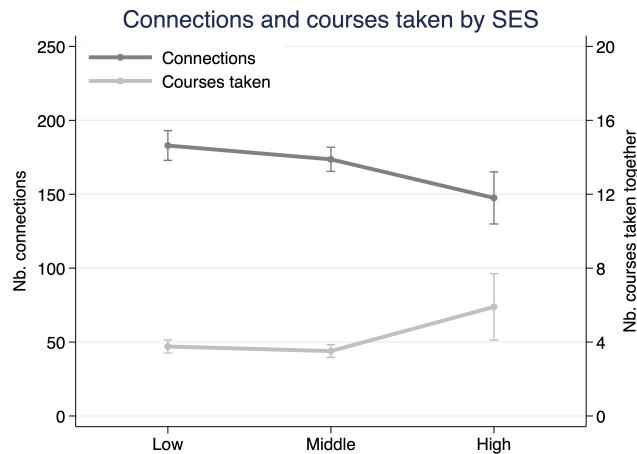
505

6.5 SES diversity in networks

505

506 How do enrollment networks differ across SES groups? We look at how the number 506
507 of connections (network size) and number of courses taken together (contact intensity) 507
508 change across SES groups in Figure 9. Low- and middle-SES students have larger net- 508
509 works but take fewer courses together with network members, while high-SES students 509
510 have smaller, denser networks. Specifically, both low- and middle-SES students have 510
511 significantly larger networks than high-SES students ($t = 3.03, p = 0.003$ and $t = 2.49,$ 511
512 $p = 0.013$, respectively), but high-SES students take significantly more courses with their 512
513 network members than both low- ($t = -3.70, p < .001$) and middle-SES ($t = -4.20,$ 513
514 $p < .001$). 514

514 Figure 9: Network size and courses taken together by SES



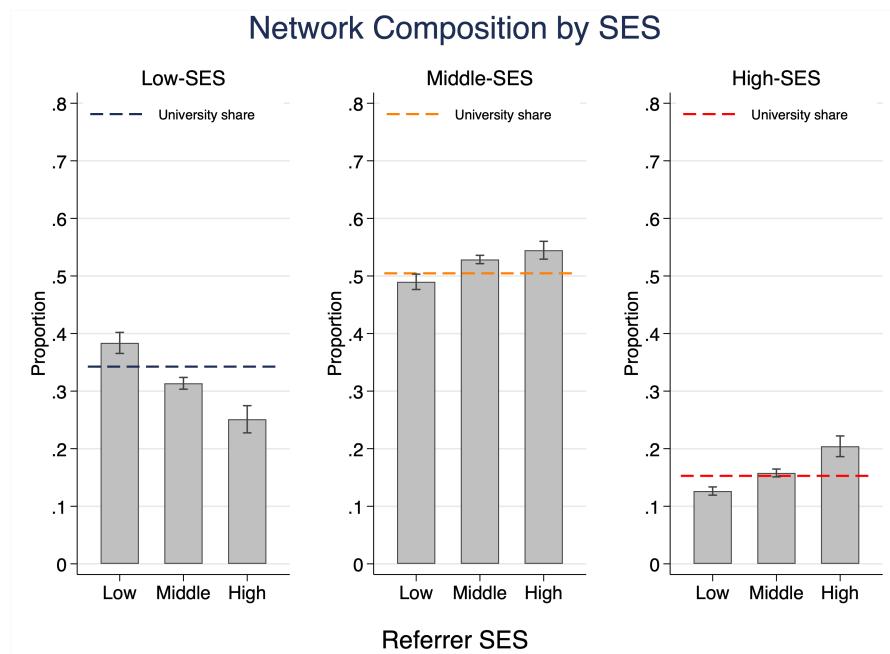
Note: This figure displays the average number of connections and the average number of classes taken together across SES groups. The data shows a decrease in the number of connections with SES, and an associated increase in the number of classes taken together.

515 What are the diversity-related consequences of SES-driven differences across net- 515
516 works? In terms of network compositions, SES groups may connect with other SES 516
517 groups at different rates than would occur randomly (Figure 10).⁸ Our results reveal 517

⁸Because we estimate the share of SES groups in every individual network, we get very precise esti-

518 modest deviations from university-wide SES composition across groups. Low-SES stu- 518
 519 dents have networks with 38.4% low-SES peers compared to the university average of 519
 520 34.3%, middle-SES students connect with 52.9% middle-SES peers versus the university 520
 521 average of 50.5%, and high-SES students show 20.4% high-SES connections compared 521
 522 to the university average of 15.3%.

Figure 10: Network shares of SES groups



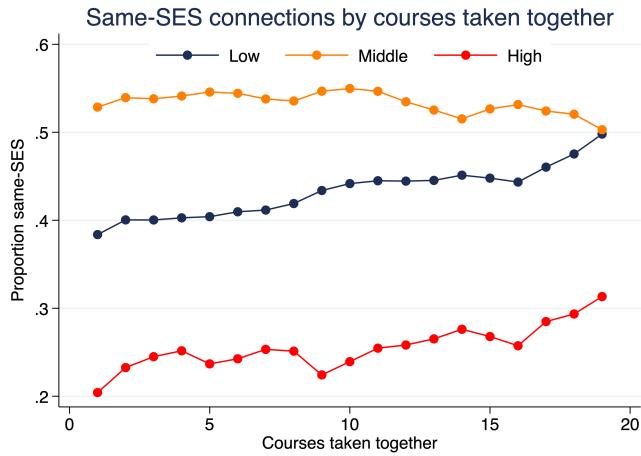
Note: This figure compares the network shares of SES groups in the networks of low-, middle-, and high-SES participants. Horizontal lines plot the university-wide shares of each SES group (Low: 35%, Mid: 51%, High: 14%). While the share of low-SES peers in the network decreases as the SES of the group increases, the share of high-SES peers in the network increases.

mates of the actual means. However, it is important to note that these are not independent observations for each network. Estimates are precise because each 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.

523 We observe larger differences between SES groups in their connection patterns with 523
524 other groups. Low-SES students connect with other low-SES students at higher rates 524
525 than middle-SES students (38.4% vs 31.4%) and high-SES students (38.4% vs 25.1%). 525
526 Conversely, high-SES students connect more with other high-SES students than both 526
527 low-SES students (20.4% vs 12.6%) and middle-SES students (20.4% vs 15.8%). Middle- 527
528 SES students are in between the two extreme patterns, connecting with middle-SES peers 528
529 at higher rates than low-SES students (52.9% vs 49.0%) but lower rates than high-SES 529
530 students (52.9% vs 54.5%). These findings indicate SES-based network segregation, with 530
531 same-SES homophily patterns across groups. 531

532 This raises an important question: What are the diversity implications of increased 532
533 connection intensity between students? As students take more courses together with 533
534 peers, the share of same-SES peers in the networks of low- and high-SES increases 534
535 while the share of middle-SES declines (see Figure 11). Both increases are substantial, 535
536 amounting to 50% for high-, and 30% for low-SES. Combining these with the earlier 536
537 result that beyond 5 courses taken together network members are almost entirely within 537
538 the same program, these suggest program selection may have strong consequences for 538
539 SES diversity in our setting. 539

Figure 11: Network size and courses taken together by courses taken

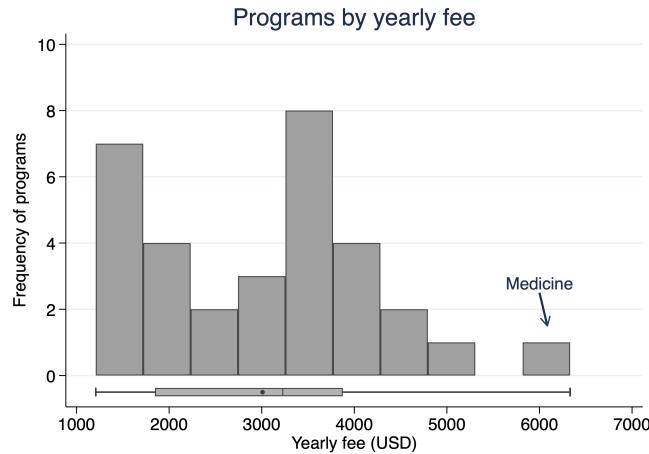


Note: This figure illustrates the shares of same-SES connections for low-, middle-, and high-SES as a function of the average number of courses taken together with network members. Low- and high-SES networks both become more homogenous as the average number of courses taken together with their connections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

540 6.6 Program selection and SES diversity 540

541 To understand the mechanisms driving these patterns, we examine program selection. 541
 542 Academic programs at this university use cost-based pricing, and typically less than 5% 542
 543 of students receive any kind of scholarship. Based on this, we first calculate how much 543
 544 every program at the university is expected to cost students per year (see Figure 12). 544
 545 Considering that net minimum monthly wage stands at \$200 and the average Colombian 545
 546 salary around \$350, the cost differences between programs are large enough to make an 546
 547 impact on program selection. Is it the case that SES groups select into programs with 547
 548 financial considerations? 548

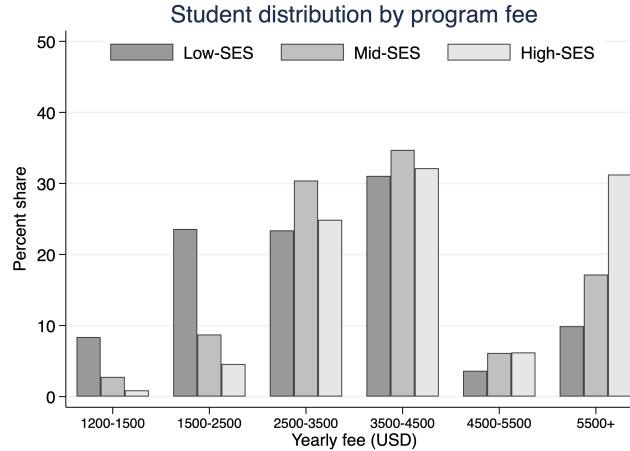
Figure 12: Programs sorted by fee



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.

We examine how SES groups are distributed across programs to identify evidence of SES-based selection (see Figure 13). 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 ($D = 31.31, p < 0.001$). With this finding, program selection could be the reason why low- and high-SES networks tend to segregate as the number of courses taken increases. The next section characterizes the referrals, and we will return to the diversity implications of program selection once we propose an understanding of how referrals were made.

Figure 13: Programs sorted by fee

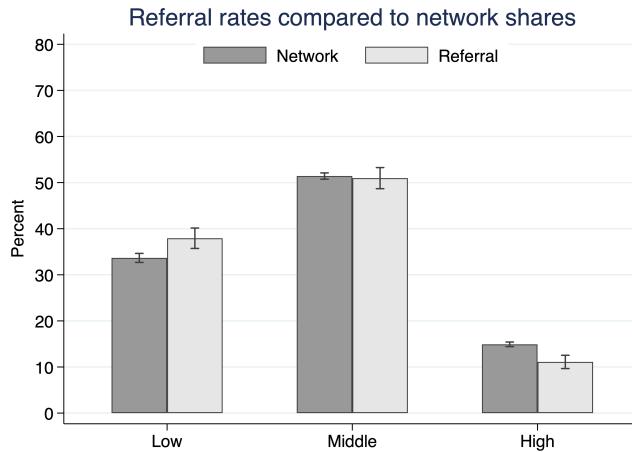


Note: This figure illustrates the distribution of each SES group across programs sorted by fee. the majority of low-SES select into programs with below average cost, while high-SES select into programs with above average cost. Medicine accounts for a third of all high-SES students at this university.

560 6.7 Referral SES composition 560

561 We now examine the overall SES compositions in referral selection. Referrals to low- 561
 562 SES peers constitute 37.9% of all referrals, compared to 33.7% low-SES representation 562
 563 in individual networks (see Figure 14). This represents a modest over-representation 563
 564 of 4.3 percentage points. For middle-SES students, referrals constitute 51.0% versus 564
 565 51.4% network representation, showing virtually no difference (-0.5 pp.). High-SES 565
 566 referrals account for 11.1% compared to 14.9% network share, an under-representation 566
 567 of 3.8 percentage points. While these patterns suggest some deviation from proportional 567
 568 representation—with slight over-referral to low-SES peers and under-referral to high-SES 568
 569 peers—the magnitudes are relatively modest. Overall, referral compositions are largely 569
 570 balanced and closely mirror the underlying network structure, with the largest deviation 570
 571 being less than 5 percentage points for any SES group. 571

Figure 14: Referral patterns compared to network composition



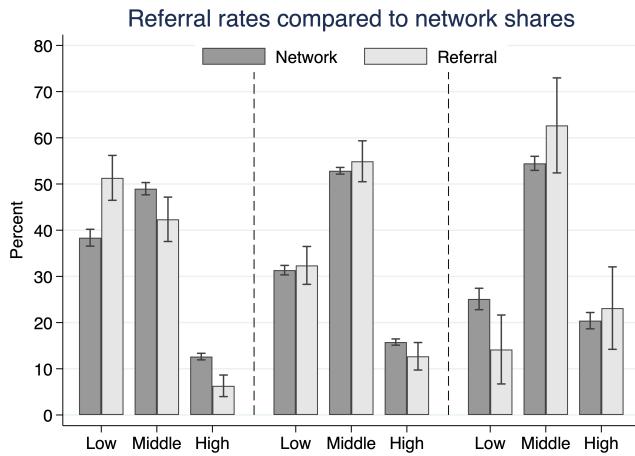
Note: This figure compares the average SES composition of referrers’ networks (dark gray) to the SES composition of referrals (light gray). Error bars represent 95% confidence intervals.

Then, we examine referral patterns by referrer SES to identify potential SES biases across groups. Figure 15 reveals mixed patterns of deviation from network composition that vary by referrer SES. Most patterns show modest deviations from network composition, with differences typically ranging from 1-6 percentage points. However, at the very extremes—low-SES to high-SES connections and vice versa—we observe the largest discrepancies between network share (which were already biased toward same-SES connections to begin with) and referral rates. Low-SES referrers show the strongest same-SES preference, referring 12.9 percentage points more to low-SES students than their network composition would suggest, while under-referring to high-SES recipients by 6.3 percentage points. Conversely, high-SES referrers under-refer to low-SES students by 10.9 percentage points compared to their network composition. Middle-SES referrers show the most balanced patterns, with deviations generally under 3 percentage points across all recipient groups. Cross-SES referral patterns, particularly between the most socioeconomically distant groups, show the largest departures from network availability. These results suggest that referral behavior diverges most from underlying network

587 structure when SES differences are most pronounced.

587

Figure 15: 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 each SES group. The panels show referral patterns for low-SES (left), middle-SES (center), and high-SES referrers (right). Error bars represent 95% confidence intervals.

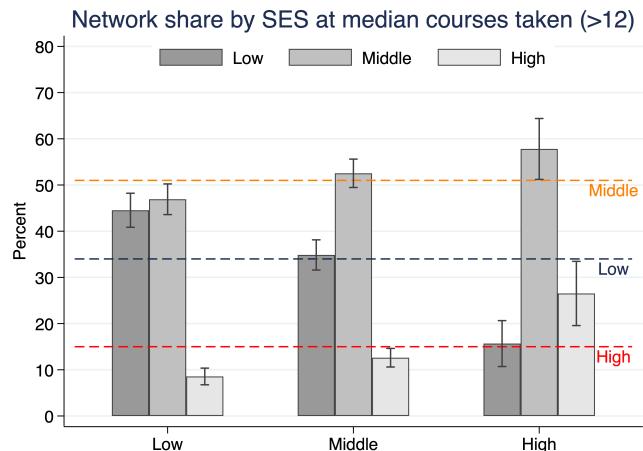
588 6.8 Ex post referral choice sets

588

589 We now shed more light on the referral behavior after having characterized how refer- 589
590 rals were made. Particularly interesting is that referrals go to peers with whom the 590
591 median participant took 12 courses, with an average of 14. By restricting the networks 591
592 for courses taken above the median, we can get a snapshot of how the referral choice 592
593 set actually looked for participants before making referral decisions. As discussed in 593
594 Section 6.5, taking more courses with network members increases the share of same-SES 594
595 individuals for both low- and high-SES students, and we had explored program selection 595
596 as a potential mechanism. In Figure 16, we show the effects of network segregation 596
597 on *ex post* referral choice sets for each SES group. Network compositions above the 597
598 median number of courses taken reveal strong segregation effects: Low-SES networks 598

599 contain 44.5% low-SES peers, higher than the 35% university-wide share by 9.5 percent- 599
 600 age points. Conversely, high-SES students are under-represented in low-SES networks at 600
 601 only 8.6% average share, compared to the 14% population share (-5.4 pp.). At the other 601
 602 extreme, high-SES networks show the reverse pattern with average low-SES share drop- 602
 603 ping to just 15.7%, a 19.3 percentage point decrease relative to the university average. 603
 604 High-SES students have a same-SES concentration at 26.5%, doubling their 14% popu- 604
 605 lation share (+12.5 pp.). Middle-SES networks remain relatively balanced and closely 605
 606 track population proportions across all SES groups. Taken together, these confirm that 606
 607 the observed referral rates of SES groups follow the network compositions above median 607
 608 number of courses taken together, except for the low-SES. We conclude that the referral 608
 609 choices in our setting are mainly driven by availability and performance. 609

Figure 16: Network size and courses taken together by courses taken



Note: This figure compares the network shares of SES groups in the networks of low-, middle-, and high-SES participants above the median number of courses taken together with peers. Horizontal lines plot the university-wide shares of each SES group (Low: 35%, Mid: 51%, High: 14%). 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.

610 **7 Robustness check**

610

611 Does the number of courses taken together have an independent effect that goes beyond 611
612 identifying peers in the same academic program? To evaluate this question we leverage 612
613 our administrative data, and identify peers within the same program: In each individ- 613
614 ual network we observe the participant-specific academic program for the participant 614
615 making the referral and alternative-specific academic program for each referral candi- 615
616 date. We add this new variable in our specification and describe our findings in Table 616
617 7. Being in the same academic program has a substantial positive effect on referral 617
618 likelihood, with coefficients ranging from 1.257 to 2.198 across all referrer SES groups. 618
619 This confirms that program affiliation serves as a strong predictor of referral decisions, 619
620 reflecting increased familiarity. Our comparison of interest is the point estimate for the 620
621 standardized number of courses taken. Across all three referrer groups, the standardized 621
622 number of courses taken together maintains its statistical significance after controlling 622
623 for same program membership. The coefficient magnitudes are expectedly smaller 623
624 compared to specifications without program controls (ranging from 0.688 to 0.930) as the 624
625 newly added variable is a moderator: Matching academic programs leads to taking more 625
626 courses together. The remaining estimates in our model remain robust to the inclusion 626
627 of the same-program variable with little change in point estimates. The persistence of 627
628 statistical significance (all $p < 0.001$) suggests that the number of courses taken together 628
629 has an independent effect on referral decisions. To sum, our measure of contact inten- 629
630 sity seems to capture meaningful social interaction patterns that lead to referrals, and 630
631 go beyond simply identifying matching academic programs.

Table 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** (0.119)	-0.140 (0.111)	-0.567* (0.331)
High-SES referral	-0.421* (0.220)	-0.249 (0.158)	-0.383 (0.281)
Entry exam (referral z-score)	0.623*** (0.054)	0.590*** (0.048)	0.892*** (0.114)
Courses taken (z-score)	0.688*** (0.032)	0.760*** (0.035)	0.930*** (0.119)
Same program	2.074*** (0.215)	2.198*** (0.185)	1.257*** (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.

632 **8 Conclusion**

632

633 In this paper, we study whether SES groups are biased toward one another beyond 633
634 what is attributable to differences in their networks, and the effects of different incentive 634
635 structures on referral behavior. Through a lab-in-the-field experiment that leverages 635
636 enrollment networks at a socially diverse university, we find that the SES biases in 636
637 referrals originate mostly from network structures, and referrals under performance-pay 637
638 incentives do not exacerbate existing SES inequalities. 638

639 Our findings reveal that enrollment networks are surprisingly segregated and referrals 639
640 from these networks reflect closely the choice sets of the referrers. We identify program 640
641 selection as the key mechanism driving this segregation. Low-SES students select into 641
642 more affordable programs, and program selection plays a major part in segregating 642
643 SES groups where low- and high-SES take more courses with their own SES group. 643
644 Consequently, referrals come almost exclusively from the same academic program as the 644
645 referrer, limiting the SES-diversity of their choice sets. Regardless of the bonus for the 645
646 referral candidate, participants also pick higher performing peers with whom they have 646
647 taken many courses together. We find that only low-SES referrers exhibit a same-SES 647
648 bias. These findings suggest that the underlying network structure plays a crucial role 648
649 in referrals, where institutional action can remedy the network segregation. 649

650 These results complement the broader literature where much of the bias in referrals 650
651 can be attributable to the “practical” choice sets of the referrers. While previous work 651
652 demonstrates that about half of referrals come from a smaller, elicited network of close 652
653 friends (Hederos et al., 2025), we go the other way and use administrative data to 653
654 construct a complete network which presumably includes close social relationships at the 654
655 institutional level. Having access to the complete network thus eliminates any potential 655
656 for under or overestimating taste-based biases (Griffith, 2022). Under performance-pay 656
657 incentives, referrers identify productive others regardless of additional financial rewards 657
658 for the referral candidate. Still, the lack of a treatment effect suggests that in both 658
659 incentive structures referrers pick close ties, shifting the responsibility to institutional 659

660 actors to create diverse environments where cross-SES social interaction can take place 660
661 more frequently and allow more diversity in networks. 661

662 These findings have policy implications. Looking forward, institutions can play a 662
663 crucial role in achieving SES equality of opportunity in higher education. Universities 663
664 are already a setting in which low-SES get exposed to typically a higher than population 664
665 share of higher-SES individuals than at other settings (Chetty et al., 2022b). Yet, 665
666 segregation within the higher education institutions remain a source for SES inequality. 666
667 If low-SES peers never get to interact in meaningful ways with higher-SES, e.g., by 667
668 taking courses together, the premise of social mobility through social channels remains 668
669 severely underexploited. Future studies should work on ways to reduce SES segregation 669
670 in collaboration with institutions, where having access to complete enrollment networks 670
671 in addition to the typical friendship elicitation methods could help identifying the exact 671
672 overlap between the two distinct approaches. 672

673 **References**

673

- 674 Alan, S., Duysak, E., Kibilay, E., & Mumcu, I. (2023). Social Exclusion and Ethnic 674
675 Segregation in Schools: The Role of Teachers' Ethnic Prejudice. *The Review of 675
676 Economics and Statistics*, 105(5), 1039–1054. doi: 10.1162/rest_a_01111 676
- 677 Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en 677
678 colombia. *Documentos CEDE*. 678
- 679 Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the 679
680 workplace: Evidence from personnel data. *Econometrica*, 77(4), 1047–1094. 680
- 681 Beaman, L., Keleher, N., & Magruder, J. (2018). Do Job Networks Disadvantage 681
682 Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor 682
683 Economics*, 36(1), 121–157. doi: 10.1086/693869 683
- 684 Beaman, L., & Magruder, J. (2012). Who Gets the Job Referral? Evidence from a 684
685 Social Networks Experiment. *American Economic Review*, 102(7), 3574–3593. 685
686 doi: 10.1257/aer.102.7.3574 686
- 687 Bolte, L., Immorlica, N., & Jackson, M. O. (2021). *The Role of Referrals in Immobility, 687
688 Inequality, and Inefficiency in Labor Markets*. (Working Paper) 688
- 689 Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory 689
690 and research for the sociology of education* (pp. 241–258). New York: Greenwood 690
691 Press. 691
- 692 Brown, M., Setren, E., & Topa, G. (2016). Do informal referrals lead to better matches? 692
693 evidence from a firm's employee referral system. *Journal of Labor Economics*, 693
694 34(1), 161–209. 694
- 695 Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on em- 695
696 ployment and inequality. *American economic review*, 94(3), 426–454. 696
- 697 Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., 697
698 ... Wernerfelt, N. (2022a). Social capital 1: Measurement and associations with 698
699 economic mobility. *Nature*, 608(7921), 108–121. doi: 10.1038/s41586-022-04996-4 699
- 700 Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... 700

- 701 Wernerfelt, N. (2022b). Social capital 2: Determinants of economic connectedness. 701
702 *Nature*, 608(7921), 122–134. doi: 10.1038/s41586-022-04997-3 702
- 703 Díaz, J., Munoz, M., Reuben, E., & Tuncer, R. (2025, March). *Peer skill identification* 703
704 *and social class: Evidence from a referral field experiment.* (Working Paper) 704
- 705 Dustmann, C., Glitz, A., Schönberg, U., & Brücker, H. (2016). Referral-based job search 705
706 networks. *The Review of Economic Studies*, 83(2), 514–546. 706
- 707 Fergusson, L., & Flórez, S. A. (2021a). Desigualdad educativa en colombia. In 707
708 J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De* 708
709 *cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* 709
710 Bogotá: Ariel. 710
- 711 Fergusson, L., & Flórez, S. A. (2021b). Distinción escolar. In J. C. Cárdenas, L. Fer- 711
712 gusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en* 712
713 *colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel. 713
- 714 Friebel, G., Heinz, M., Hoffman, M., & Zubanov, N. (2023). What do employee referral 714
715 programs do? measuring the direct and overall effects of a management practice. 715
716 *Journal of Political Economy*, 131(3), 633–686. 716
- 717 García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la 717
718 cuna: La movilidad social a través de la educación en los municipios de colombia. 718
719 *Documentos CEDE*. 719
- 720 García Villegas, M., & Cobo, P. (2021). La dimensión cultural del apartheid educativo. 720
721 In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De* 721
722 *cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* 722
723 Bogotá: Ariel. 723
- 724 Gee, L. K., Jones, J. J., & Burke, M. (2017). Social networks and labor markets: How 724
725 strong ties relate to job finding on facebook's social network. *Journal of Labor* 725
726 *Economics*, 35(2), 485–518. 726
- 727 Griffith, A. (2022). Name Your Friends, but Only Five? The Importance of Censoring in 727
728 Peer Effects Estimates Using Social Network Data. *Journal of Labor Economics*. 728
729 doi: 10.1086/717935 729

- 730 Guevara S, J. D., & Shields, R. (2019). Spatializing stratification: Bogotá. *Ardeth. A* 730
731 *Magazine on the Power of the Project*(4), 223–236. 731
- 732 Hederos, K., Sandberg, A., Kvissberg, L., & Polano, E. (2025). Gender homophily 732
733 in job referrals: Evidence from a field study among university students. *Labour* 733
734 *Economics*, 92, 102662. 734
- 735 Hudson, R. A., & Library of Congress (Eds.). (2010). *Colombia: a country study* 735
736 (5th ed.). Washington, D.C: Federal Research Division, Library of Congress: For 736
737 sale by the Supt. of Docs., U.S. G.P.O. Retrieved from the Library of Congress, 737
738 <https://www.loc.gov/item/2010009203/>. 738
- 739 Jaramillo-Echeverri, J., & Álvarez, A. (2023). *The Persistence of Segregation in Edu* 739
740 *cation: Evidence from Historical Elites and Ethnic Surnames in Colombia* (SSRN 740
741 Scholarly Paper No. 4575894). Rochester, NY: Social Science Research Network. 741
742 doi: 10.2139/ssrn.4575894 742
- 743 Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. 743
744 *American Journal of Sociology*, 115(2), 405–450. Retrieved from <https://www> 744
745 [.journals.uchicago.edu/doi/abs/10.1086/599247](https://journals.uchicago.edu/doi/abs/10.1086/599247) doi: 10.1086/599247 745
- 746 Kramarz, F., & Nordström Skans, O. (2014). When strong ties are strong: Networks and 746
747 youth labour market entry. *The Review of Economic Studies*, 81(3), 1164–1200. 747
- 748 Lin, N., Ensel, W. M., & Vaughn, J. C. (1981). Social Resources and Strength of 748
749 Ties: Structural Factors in Occupational Status Attainment. *American Sociological 749
750 Review*, 46(4), 393–405. doi: 10.2307/2095260 750
- 751 Loury, G. C. (1977). A dynamic theory of racial income differences. In P. A. Wallace 751
752 & A. M. LaMond (Eds.), *Women, minorities, and employment discrimination* 752
753 (pp. 153–186). Lexington, MA: Lexington Books. (Originally published as Dis- 753
754 cussion Paper 225, Northwestern University, Center for Mathematical Studies in 754
755 Economics and Management Science, 1976) 755
- 756 McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily 756
757 in social networks. *Annual review of sociology*, 27(1), 415–444. 757
- 758 Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an 758

- 759 Economic Analysis. *American Economic Review*. 759

760 Mouw, T. (2003). Social Capital and Finding a Job: Do Contacts Matter? *American 760
Sociological Review*, 68(6), 868–898. doi: 10.1177/000312240306800604 761

761 762 Pallais, A., & Sands, E. G. (2016). Why the Referential Treatment? Evidence from 762
Field Experiments on Referrals. *Journal of Political Economy*, 124(6), 1793–1828. 763

763 764 doi: 10.1086/688850 764

765 Pedulla, D. S., & Pager, D. (2019). Race and networks in the job search process. 765
American Sociological Review, 84, 983–1012. doi: 10.1177/0003122419883255 766

766 767 Rohrer, J. M., Keller, T., & Elwert, F. (2021). Proximity can induce diverse friendships: 767
A large randomized classroom experiment. *PLOS ONE*, 16(8), e0255097. doi: 768

768 769 10.1371/journal.pone.0255097 769

770 Smith, S. S. (2005). “Don’t put my name on it”: Social Capital Activation and Job- 770
Finding Assistance among the Black Urban Poor. *American Journal of Sociology*, 771
111(1), 1–57. doi: 10.1086/428814 772

772 773 Stansbury, A., & Rodriguez, K. (2024). The class gap in career progression: Evidence 773
from US academia. *Working Paper*. 774

774 775 Topa, G. (2019). Social and spatial networks in labour markets. *Oxford Review of 775
Economic Policy*, 35(4), 722–745. 776

776 777 United Nations. (2023). *Social panorama of latin america and the caribbean 777
2023: labour inclusion as a key axis of inclusive social development*. 778
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> 779

779 780 Uribe-Mallarino, C. (2008). Estratificación social en bogotá: de la política pública a la 780
dinámica de la segregación social. *Universitas humanistica*(65), 139–172. 781

781 782 Wang, S.-Y. (2013). Marriage networks, nepotism, and labor market outcomes in china. 782
American Economic Journal: Applied Economics, 5(3), 91–112. 783

783 784 Witte, M. (2021). Why do workers make job referrals? experimental evidence from 784
ethiopia. *Working Paper*. 785

785 786 787 788

788 World Bank. (2024). *Regional poverty and inequality update spring 2024* 788
789 (Poverty and Equity Global Practice Brief). Washington, D.C.: World 789
790 Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269> 790
791 791

792 **A Additional Figures and Tables**

792

793 **Additional Figures**

793

Table A.1: Selection into the experiment

	University	Sample	<i>p</i>
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	<i>p</i>
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

Referrer SES	Network average for SES group		
	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 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.

795 **B Experiment**

795

796 We include the English version of the instructions used in Qualtrics. Participants saw 796
797 the Spanish version. Horizontal lines in the text indicate page breaks and clarifying 797
798 comments are inside brackets. 798

799 **Consent**

799

800 You have been invited to participate in this decision-making study. This study is directed 800
801 by [omitted for anonymous review] and organized with the support of the Social Bee Lab 801
802 (Social Behavior and Experimental Economics Laboratory) at UNAB. 802

803 In this study, we will pay **one (1)** out of every **ten (10)** participants, who will be 803
804 randomly selected. Each selected person will receive a fixed payment of **70,000** (seventy 804
805 thousand pesos) for completing the study. Additionally, they can earn up to **270,000** 805
806 (two hundred and seventy thousand pesos), depending on their decisions. So, in total, 806
807 if you are selected to receive payment, you can earn up to **340,000** (three hundred and 807
808 forty thousand pesos) for completing this study. 808

809 If you are selected, you can claim your payment at any Banco de Bogotá office by 809
810 presenting your ID. Your participation in this study is voluntary and you can leave the 810
811 study at any time. If you withdraw before completing the study, you will not receive 811
812 any payment. 812

813 The estimated duration of this study is 20 minutes. 813

814 The purpose of this study is to understand how people make decisions. For this, we will 814
815 use administrative information from the university such as the SABER 11 test scores of 815
816 various students (including you). Your responses will not be shared with anyone and your 816
817 participation will not affect your academic records. To maintain strict confidentiality, the 817
818 research results will not be associated at any time with information that could personally 818

819 identify you. 819

820 There are no risks associated with your participation in this study beyond everyday risks. 820
821 However, if you wish to report any problems, you can contact Professor [omitted for 821
822 anonymous review]. For questions related to your rights as a research study participant, 822
823 you can contact the IRB office of [omitted for anonymous review]. 823

824 By selecting the option “I want to participate in the study” below, you give your con- 824
825 sent to participate in this study and allow us to compare your responses with some 825
826 administrative records from the university. 826

- 827 • I want to participate in the study [advances to next page] 827
828 • I do not want to participate in the study 828

829 _____ 829

830 Student Information 830

831 Please write your student code. In case you are enrolled in more than one program 831
832 simultaneously, write the code of the first program you entered: 832

833 [Student ID code] 833

834 What semester are you currently in? 834

835 [Slider ranging from 1 to 11] 835

836 _____ 836

837 [Random assignment to treatment or control] 837

838 **Instructions**

838

839 The instructions for this study are presented in the following video. Please watch it 839
840 carefully. We will explain your participation and how earnings are determined if you are 840
841 selected to receive payment. 841

842 [Treatment-specific instructions in video format] 842

843 If you want to read the text of the instructions narrated in the video, press the “Read 843
844 instruction text” button. Also know that in each question, there will be a button with 844
845 information that will remind you if that question has earnings and how it is calculated, 845
846 in case you have any doubts. 846

847 • I want to read the instructions text [text version below] 847

848 —————— 848

849 In this study, you will respond to three types of questions. First, are the belief questions. 849
850 For belief questions, we will use as reference the results of the SABER 11 test that you 850
851 and other students took to enter the university, focused on three areas of the exam: 851
852 mathematics, reading, and English. 852

853 For each area, we will take the scores of all university students and order them from 853
854 lowest to highest. We will then group them into 100 percentiles. The percentile is a 854
855 position measure that indicates the percentage of students with an exam score that is 855
856 above or below a value. 856

857 For example, if your score in mathematics is in the 20th percentile, it means that 20 857
858 percent of university students have a score lower than yours and the remaining 80 percent 858
859 have a higher score. A sample belief question is: “compared to university students, in 859
860 what percentile is your score for mathematics?” 860

861 If your answer is correct, you can earn 20 thousand pesos. We say your answer is correct 861

862 if the difference between the percentile you suggest and the actual percentile of your 862
863 score is not greater than 7 units. For example, if you have a score that is in the 33rd 863
864 percentile and you say it is in the 38th, the answer is correct because the difference is 864
865 less than 7. But if you answer that it is in the 41st, the difference is greater than 7 and 865
866 the answer is incorrect. 866

867 The second type of questions are recommendation questions and are also based on the 867
868 mathematics, reading, and English areas of the SABER 11 test. We will ask you to think 868
869 about the students with whom you have taken or are taking classes, to recommend from 869
870 among them the person you consider best at solving problems similar to those on the 870
871 SABER 11 test. 871

872 When you start typing the name of your recommended person, the computer will show 872
873 suggestions with the full name, program, and university entry year of different students. 873
874 Choose the person you want to recommend. If the name doesn't appear, check that you 874
875 are writing it correctly. Do not use accents and use 'n' instead of 'ñ'. If it still doesn't 875
876 appear, it may be because that person is not enrolled this semester or because they did 876
877 not take the SABER 11 test. In that case, recommend someone else. 877

878 You can earn up to 250,000 pesos for your recommendation. We will multiply your 878
879 recommended person's score by 100 pesos if they are in the first 50 percentiles. We will 879
880 multiply it by 500 pesos if your recommended person's score is between the 51st and 880
881 65th percentile. If it is between the 66th and 80th percentile, we will multiply your 881
882 recommended person's score by 1000 pesos. If the score is between the 81st and 90th 882
883 percentile, you earn 1500 pesos multiplied by your recommended person's score. And if 883
884 the score is between the 91st and 100th percentile, we will multiply your recommended 884
885 person's score by 2500 pesos to determine the earnings. 885

886 The third type of questions are information questions and focus on aspects of your 886
887 personal life or your relationship with the people you have recommended. 887

888 **Earnings** 888

889 Now we will explain who gets paid for participating and how the earnings for this study 889
890 are assigned. The computer will randomly select one out of every 10 participants to pay 890
891 for their responses. For selected individuals, the computer will randomly choose one of 891
892 the three areas, and from that chosen area, it will pay for one of the belief questions. 892

893 Similarly, the computer will randomly select one of the three areas to pay for one of the 893
894 recommendation questions. 894

895 **Additionally, if you are selected to receive payment, your recommended per- 895
896 son in the chosen area will receive a fixed payment of 100 thousand pesos. 896
897 [Only seen if assigned to the treatment]** 897

898 Each person selected to receive payment for this study can earn: up to 20 thousand pesos 898
899 for one of the belief questions, up to 250 thousand pesos for one of the recommendation 899
900 questions, and a fixed payment of 70 thousand pesos for completing the study. 900

901 Selected individuals can earn up to 340 thousand pesos. 901

902 902

903 [Participants go through all three Subject Areas in randomized order] 903

904 **Subject Areas** 904

905 **Critical Reading** 905

906 For this section, we will use as reference the Critical Reading test from SABER 11, which 906
907 evaluates the necessary competencies to understand, interpret, and evaluate texts that 907
908 can be found in everyday life and in non-specialized academic fields. 908

909 [Clicking shows the example question from SABER 11 below] 909

910 Although the democratic political tradition dates back to ancient Greece, political 910
911 thinkers did not address the democratic cause until the 19th century. Until then, democ- 911
912 racy had been rejected as the government of the ignorant and unenlightened masses. 912
913 Today it seems that we have all become democrats without having solid arguments in 913
914 favor. Liberals, conservatives, socialists, communists, anarchists, and even fascists have 914
915 rushed to proclaim the virtues of democracy and to show their democratic credentials 915
916 (Andrew Heywood). According to the text, which political positions identify themselves 916
917 as democratic? 917

- 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

923 Mathematics 923

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.

926 [Clicking shows the example question from SABER 11 below] 926

927 A person living in Colombia has investments in dollars in the United States and knows 927
928 that the exchange rate of the dollar against the Colombian peso will remain constant 928
929 this month, with 1 dollar equivalent to 2,000 Colombian pesos. Their investment, in 929
930 dollars, will yield profits of 3% in the same period. A friend assures them that their 930
931 profits in pesos will also be 3%. Their friend's statement is: 931

- | | | |
|-----|---|-----|
| 934 | <ul style="list-style-type: none">• Incorrect. The exact value of the investment should be known. | 934 |
| 935 | <ul style="list-style-type: none">• Correct. 3% is a fixed proportion in either currency. | 935 |
| 936 | <ul style="list-style-type: none">• Incorrect. 3% is a larger increase in Colombian pesos. | 936 |

937 937

938 English 938

939 This section uses the English test from SABER 11 as a reference, which evaluates that 939
940 the person demonstrates their communicative abilities in reading and language use in 940
941 this language. 941

⁹⁴² [Clicking shows the example question from SABER 11 below]

943 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] 953

954 **Your Score**

954

955 Compared to university students, in which percentile do you think your [Subject Area] 955
956 test score falls (1 is the lowest percentile and 100 the highest)? 956

957 [Clicking shows the explanations below] 957

958 How is a percentile calculated? 958

959 A percentile is a position measurement. To calculate it, we take the test scores for all 959
960 students currently enrolled in the university and order them from lowest to highest. The 960
961 percentile value you choose refers to the percentage of students whose score is below 961
962 yours. For example, if you choose the 20th percentile, you're indicating that 20% of 962
963 students have a score lower than yours and the remaining 80% have a score higher than 963
964 yours. 964

965 What can I earn for this question? 965

966 For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the 966
967 difference between your response and the correct percentile is less than 7. For example, if 967
968 the percentile where your score falls is 33 and you respond with 38 (or 28), the difference 968
969 is 5 and the answer is considered correct. But if you respond with 41 or more (or 25 or 969
970 less), for example, the difference would be greater than 7 and the answer is incorrect. 970

971 Please move the sphere to indicate which percentile you think your score falls in: 971

972 [Slider with values from 0 to 100] 972

973

 973

974 **Recommendation**

974

975 Among the people with whom you have taken any class at the university, who is your 975
976 recommendation for the [Subject Area] test? Please write that person's name in the 976
977 box below: 977

978 **Important:** You will not be considered for payment unless the recommended 978
979 person is someone with whom you have taken at least one class during your 979
980 studies. 980

981 Your response is only a recommendation for the purposes of this study and we will **not** 981
982 contact your recommended person at any time. 982

983 [Clicking shows the explanations below] 983

984 Who can I recommend? 984

985 Your recommendation **must** be someone with whom you have taken (or are taking) a 985
986 class. If not, your answer will not be considered for payment. The person you recommend 986
987 will not be contacted or receive any benefit from your recommendation. 987

988 As you write, you will see up to 7 suggested student names containing the letters you 988
989 have entered. The more you write, the more accurate the suggestions will be. Please 989
990 write **without** accents and use the letter 'n' instead of 'ñ'. If the name of the person 990
991 you're writing doesn't appear, it could be because you made an error while writing the 991
992 name. 992

993 If the name is correct and still doesn't appear, it could be because the student is not en- 993
994 rolled this semester or didn't take the SABER 11 test. In that case, you must recommend 994
995 someone else. 995

996 My earnings for this question? 996

997 For your recommendation, you could receive earnings of up to 250,000 (two hundred and 997
998 fifty thousand) PESOS. The earnings are calculated based on your recommendation's 998
999 score and the percentile of that score compared to other UNAB students, as follows: 999

- 1000 • We will multiply your recommendation's score by \$100 (one hundred) pesos if it's 1000
1001 between the 1st and 50th percentiles 1001
- 1002 • We will multiply your recommendation's score by \$500 (five hundred) pesos if it's 1002
1003 between the 51st and 65th percentiles 1003
- 1004 • We will multiply your recommendation's score by \$1000 (one thousand) pesos if 1004
1005 it's between the 66th and 80th percentiles 1005
- 1006 • We will multiply your recommendation's score by \$1500 (one thousand five hun- 1006
1007 dred) pesos if it's between the 81st and 90th percentiles 1007
- 1008 • We will multiply your recommendation's score by \$2500 (two thousand five hun- 1008
1009 dred) pesos if it's between the 91st and 100th percentiles 1009

1010 This is illustrated in the image below: 1010

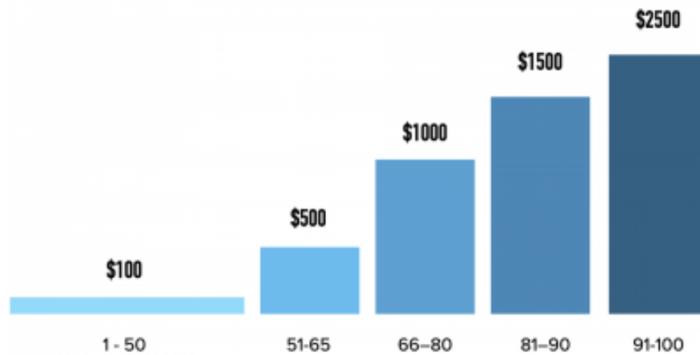


Figure B.1: Earnings for recommendation questions

1011 For example, if your recommendation got 54 points and the score is in the 48th percentile, 1011

1012 you could earn $54 \times 100 = 5400$ PESOS. But, if the same score of 54 points were in the 1012
1013 98th percentile, you could earn $54 \times 2500 = 135,000$ PESOS. 1013

1014 [Text field with student name suggestions popping up as participant types] 1014

1015 _____ 1015

1016 Relationship with your recommendation 1016

1017 How close is your relationship with your recommendedation: “[Name of the student 1017
1018 selected from earlier]”? (0 indicates you are barely acquaintances and 10 means you are 1018
1019 very close) 1019

1020 [Slider with values from 0 to 10] 1020

1021 _____ 1021

1022 Your recommendation's score 1022

1023 Compared to university students, in which percentile do you think [Name of the student 1023
1024 selected from earlier]’s score falls in the [Subject Area] test (1 is the lowest percentile 1024
1025 and 100 the highest)? 1025

1026 [Clicking shows the explanations below] 1026

1027 How is a percentile calculated? 1027

1028 A percentile is a position measurement. To calculate it, we take the test scores for all 1028
1029 students currently enrolled in the university and order them from lowest to highest. The 1029
1030 percentile value you choose refers to the percentage of students whose score is below 1030
1031 yours. For example, if you choose the 20th percentile, you’re indicating that 20% of 1031
1032 students have a score lower than yours and the remaining 80% have a score higher than 1032
1033 yours. 1033

1034 What can I earn for this question?

1034

1035 For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the 1035
1036 difference between your response and the correct percentile is less than 7. For example, 1036
1037 if the percentile where your recommended person's score falls is 33 and you respond with 1037
1038 38 (or 28), the difference is 5 and the answer is considered correct. But if you respond 1038
1039 with 41 or more (or 25 or less), for example, the difference would be greater than 7 and 1039
1040 the answer is incorrect.

1040

1041 Please move the sphere to indicate which percentile you think your recommended per- 1041
1042 son's score falls in:

1042

1043 [Slider with values from 0 to 100]

1043

1044 _____ 1044

1045 Demographic Information

1045

1046 What is the highest level of education achieved by your father?

1046

1047 [Primary, High School, University, Graduate Studies, Not Applicable]

1047

1048 What is the highest level of education achieved by your mother?

1048

1049 [Primary, High School, University, Graduate Studies, Not Applicable]

1049

1050 Please indicate the socio-economic group to which your family belongs:

1050

1051 [Group A (Strata 1 or 2), Group B (Strata 3 or 4), Group C (Strata 5 or 6)]

1051

1052 _____ 1052

1053 UNAB Students Distribution

1053

1054 Thinking about UNAB students, in your opinion, what percentage belongs to each socio- 1054
1055 economic group? The total must sum to 100%: 1055

1056 [Group A (Strata 1 or 2) percentage input area] 1056
1057 [Group B (Strata 3 or 4) percentage input area] 1057
1058 [Group C (Strata 5 or 6) percentage input area] 1058
1059 [Shows sum of above percentages] 1059

1060 ————— 1060

1061 End of the Experiment

1061

1062 Thank you for participating in this study. 1062

1063 If you are chosen to receive payment for your participation, you will receive a confirma- 1063
1064 tion to your UNAB email and a link to fill out a form with your information. The process 1064
1065 of processing payments is done through Nequi and takes approximately 15 business days, 1065
1066 counted from the day of your participation. 1066

1067 [Clicking shows the explanations below] 1067

1068 Who gets paid and how is it decided? 1068

1069 The computer will randomly select one out of every ten participants in this study to be 1069
1070 paid for their decisions. 1070

1071 For selected individuals, the computer will randomly select one area: mathematics, 1071
1072 reading, or English, and from that area will select one of the belief questions. If the 1072
1073 answer to that question is correct, the participant will receive 20,000 pesos. 1073

1074 The computer will randomly select an area (mathematics, critical reading, or English) to 1074
1075 pay for one of the recommendation questions. The area chosen for the recommendation 1075
1076 question is independent of the area chosen for the belief question. The computer will 1076
1077 take one of the two recommendations you have made for the chosen area. Depending on 1077
1078 your recommendation's score, you could win up to 250,000 pesos. 1078

1079 Additionally, people selected to receive payment for their participation will have a fixed 1079
1080 earnings of 70,000 pesos for completing the study. 1080

1081 _____ 1081

1082 **Participation** 1082

1083 In the future, we will conduct studies similar to this one where people can earn money 1083
1084 for their participation. The participation in these studies is by invitation only. Please 1084
1085 indicate if you are interested in being invited to other studies similar to this one: 1085

1086 [Yes, No] 1086