

Class differences in social networks: Evidence from a referral experiment

Manuel Munoz*, Ernesto Reuben^{†*}, Reha Tuncer[‡]

August 1, 2025

Abstract

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

*Luxembourg Institute of Socio-Economic Research

[†]Division of Social Science, New York University Abu Dhabi

[‡]University of Luxembourg

20 the university share. Yet, we find no bias against low-SES individuals once we account 20
21 for network structures. We randomly assign half of the participants to a condition where 21
22 their referral candidate receives a fixed bonus on top of pay-for-performance referral 22
23 incentives. We find that additional incentives for the referral candidate do not change 23
24 connection intensity with the referral nor the referral quality. Our findings suggest that 24
25 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 empirical strategy and 184
185 in Section 4 we present the design of the experiment. In Section 5 we describe the 185
186 experimental sample, incentives and the procedure. Section 6 discusses the results of 186
187 the experiment and Section 7 discusses potential mechanisms and robustness checks. 187
188 Section 8 concludes. The Appendix presents additional tables and figures as well as the 188
189 experiment instructions. 189

190 **2 Background and Setting**

190

191 **2.1 Inequality and SES in Colombia**

191

192 Our experiment took place in Colombia, a country that consistently ranks highly in 192
193 terms of economic inequality. The richest decile of Colombians earn 50 times more than 193
194 the poorest decile ([United Nations, 2023](#); [World Bank, 2024](#)). This economic disparity 194
195 creates profound differences in outcomes across SES groups in terms of education, geo- 195
196 geographic residence, language, manners, and social networks ([Angulo et al., 2012](#); [García 196
197 et al., 2015](#); [García Villegas & Cobo, 2021](#)). While these patterns are not atypical and 197
198 exist elsewhere, Colombia's pronounced inequality makes economic, educational, and 198
199 cultural differences across SES particularly visible.

199

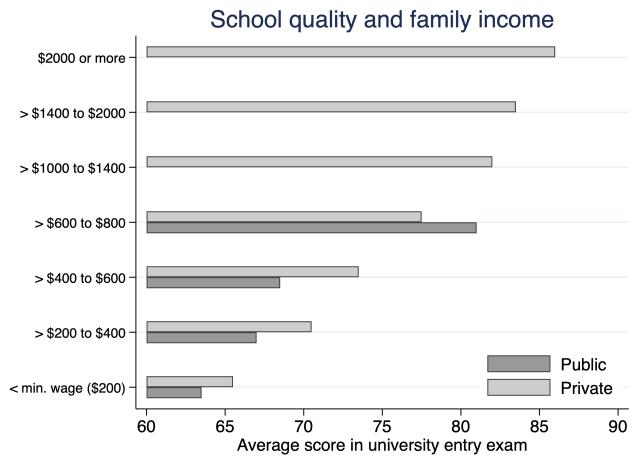
200 We rely on Colombia's established estrato classification system to measure SES in 200
201 our study. In 1994, Colombia introduced a nationwide system that divides the popula- 201
202 tion into six strata based on "similar social and economic characteristics" ([Hudson & 202
203 Library of Congress, 2010](#), p. 102). Designed for utility subsidies from higher strata to 203
204 support lower strata, the system aligns with and reinforces existing social class divisions 204
205 ([Guevara S & Shields, 2019](#); [Uribe-Mallarino, 2008](#)). It is widely used by policymakers 205
206 and in official statistics ([Fergusson & Flórez, 2021a](#)). Using the estrato system, we cat- 206
207 egorize students in strata 1-2 as low-SES, strata 3-4 as middle-SES, and strata 5-6 as 207
208 high-SES.

208

209 In higher education, Colombia's pronounced economic equality manifests itself by 209
210 preventing meaningful interaction between SES groups. Wealthy families attend ex- 210
211 clusive private schools while poorer families access lower-quality public or "non-elite" 211
212 private institutions (see Figure 1). Taken together, the unique ways in which economic 212
213 inequality manifests itself in the Colombian higher educational setting provides the ideal 213
214 conditions to study biases related to SES in referral selection.

214

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 scores in the 65-70 band (see Appendix Figures A.1a and A.1b), the private university where we conducted this experiment caters to low-, middle- and high-SES students. Figure reproduced from [Fergusson and Flórez \(2021b\)](#).

2.2 Partner institution and the enrollment network

Our study takes place in a non-elite private university which attracts students across the socioeconomic spectrum: The university's undergraduate student body comprises 35% low-SES, 50% middle-SES, and 15% high-SES students.⁶ This diversity provides opportunities for different SES groups to meet and interact within the same institutional framework.

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

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

227 Undergraduate students at the university choose among 32 different academic pro- 227
228 grams. Students take between 5 to 7 courses per semester, and programs last anywhere 228
229 between 4 to 12 semesters (2 to 6 years). The majority (64%) of students are enrolled 229
230 in the 10 programs described in Appendix Figure A.2. Medicine, the largest program 230
231 by size at the university, lasts for 12 semesters, followed by engineering programs at 10 231
232 semesters. Most remaining programs last for about 8 to 10 semesters, with specialized 232
233 programs for immediate entry into the workforce lasting only 4 semesters. 233

234 To map social connections at the university, we construct enrollment networks using 234
235 administrative data. For each participant, we identify all other undergraduate students 235
236 with whom they have taken at least one course and create their individual network of 236
237 university connections. The size of this network depends on how many different students 237
238 a participant has encountered through coursework, while the intensity of connection is 238
239 measured by the number of courses taken together. This approach provides a complete 239
240 picture of each participant's social environment at the university, including detailed 240
241 characteristics (i.e., SES, academic program, performance) for both the participant and 241
242 every person in their network. 242

243 3 Empirical Strategy 243

244 We use a conditional logit model to study SES biases in referral selection. Our depen- 244
245 dent variable follows a multinomial distribution where referrer i selects one candidate 245
246 j from their enrollment network for two exam areas. For each referrer, we observe all 246
247 potential candidates, i.e, students they took at least one course with, along with their 247
248 characteristics. The conditional logit model with individual fixed effects takes the form: 248

$$249 Y_{ij} = \alpha_i + \beta_1 SES_{ij} + \beta_2 X_{ij} + \varepsilon_{ij} \quad (1)$$

250 where $Y_{ij} = 1$ if referrer i selects candidate j , and 0 otherwise. We set middle-SES 249
251 as the base category, so β_1 is the log-odds estimate for referring low- and high-SES 250
251 candidates relative to middle-SES. X_{ij} includes the remaining characteristics of referral 251

252 candidates in the enrollment network that improve model fit such as entry exams scores 252
253 and the number of courses taken together with the referrer. These continuous variables 253
254 are standardized using means and standard deviations calculated by first computing 254
255 network-level statistics for each referrer, then averaging across all 734 networks.⁷ The 255
256 individual fixed effects α_i control for all referrer-specific factors that might influence 256
257 both network formation and referral decisions. Because we observe two referrals from 257
258 each referrer, we cluster standard errors at the referrer level to account for the potential 258
259 correlation within these referral decisions. 259

260 The key advantage of this approach is that by conditioning on each referrer's enroll- 260
261 ment network, we eliminate selection bias from program choice and other factors that 261
262 determine who appears in each person's choice set. The identifying variation comes from 262
263 within-network differences in referral decisions, holding constant the pool of available 263
264 candidates. 264

265 We estimate separate models for each referrer SES group to estimate aggregate SES 265
266 biases across socioeconomic groups. 266

267 For causal identification, we require two assumptions. First, conditional exogeneity. 267
268 SES and the number of courses taken together could be endogenous due to program se- 268
269 lection. High-SES students sort into expensive programs while low-SES students choose 269
270 affordable programs, creating SES variation across enrollment networks. Similarly, the 270
271 number of courses taken together reflects program selection decisions that may correlate 271
272 with unobserved referral preferences. However, conditional on the realized enrollment 272
273 network, the remaining variation in both SES and the number of courses taken together 273
274 across referral candidates must be independent of unobserved factors affecting referral 274
275 decisions. As a robustness check, we show that being in the same program with the refer- 275
276 rer does not impact our SES bias estimates, although it reduces the coefficient estimate 276
277 for the number of courses taken together. 277

⁷Each referral candidate's entry exam score and the number of courses they have taken with the referrer is standardized using these sample-level statistics. The standardization formula is $z_i = (x_i - \bar{X})/\sigma$, where \bar{X} and σ are the average mean and standard deviation across participant networks for the measure.

278 Second, the independence of irrelevant alternatives. This assumption could be vio- 278
279 lated if peers within the same SES group are viewed as close substitutes, where adding 279
280 similar alternatives distorts choice probabilities. While this concern may have some 280
281 validity in our setting,⁸ Alternative discrete choice models that relax IIA are computa- 281
282 tionally prohibitive given our large dataset.⁹ We therefore proceed with the conditional 282
283 logit framework while acknowledging this limitation. 283

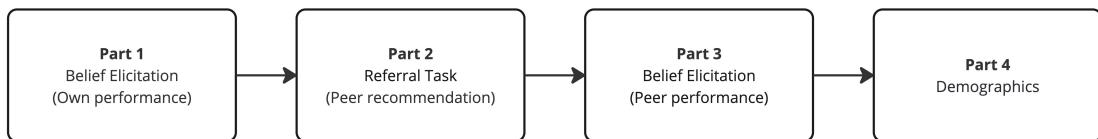
284 4 Design 284

285 We designed an online experiment to assess peer referral selection from an SES perspec- 285
286 tive and to evaluate the causal effect of providing a bonus to referral candidates. The 286
287 experimental design consisted of two incentivized tasks administered in the following 287
288 sequence: First, participants completed belief elicitation tasks about their own perfor- 288
289 mance on the national university entry exam. Second, they completed the main referral 289
290 task, nominating peers based on exam performance in two academic areas. Finally, par- 290
291 ticipants reported beliefs about their referrals' performance and provided demographic 291
292 information. This structure allowed us to measure both the accuracy of participants' 292
293 beliefs and their referral behavior under controlled incentive conditions. Figure 2 shows 293
294 the experimental timeline, and detailed instructions are provided in Appendix B. 294

⁸Among participants making referrals to two different individuals, half refer to someone else from the same SES, suggesting potential substitutability within SES groups.

⁹Models such as nested logit become computationally intractable with over 250,000 observations across 734 individuals.

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.

295 4.1 Performance measures

296 To establish an objective basis for referral performance, we use national university entry 296
297 exam scores (SABER 11). All Colombian high school students take the SABER 11 297
298 exam at the end of their final year as a requirement for university admission. The scores 298
299 from this exam provide pre-existing, comparable measures of performance. By using 299
300 existing administrative data, we also ensure that all eligible students have comparable 300
301 performance measures. 301

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

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

313 **4.2 Referral task**

313

314 The main task involves making referrals among peers. For both exam areas (critical 314
315 reading and mathematics), participants refer one peer they believe excels in that area. 315
316 We provide an example question from the relevant exam area to clarify the skills that 316
317 are being assessed. Participants type the name of their preferred candidate to make 317
318 a referral. To avoid issues with recall, the interface provides autocomplete name and 318
319 program suggestions from the administrative database (see Figure 3). 319

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:

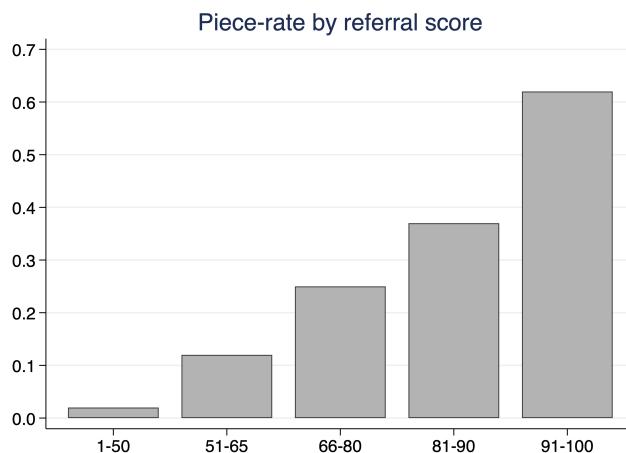
John
John Lennon (Music - 2018) 
John Stuart Mill (Law - 2020)

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

320 Participants can only refer students with whom they have taken at least one class 320
321 during their university studies. This requirement ensures that referrals are based on 321
322 actual peer interactions. We randomize the order in which participants make referrals 322
323 across the two exam areas. 323

324 We incentivize referrals using a piece rate payment structure. Referrers earn in- 324
325 creasing payments as the percentile ranking of their referral increases (see Figure 4). We 325
326 multiply the piece rate coefficient associated with the percentile rank by the actual exam 326
327 scores of the referral to calculate earnings. This payment structure provides strong in- 327
328 centives to refer highly ranked peers with potential earnings going up to \$60 per referral. 328

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.

330 **4.3 Bonus Treatment**

331 To examine how different incentive structures affect referral selection, we randomly assign 331
 332 a fixed bonus payment for students who get a referral. In the **Baseline** treatment, only 332
 333 the participants can earn money based on their referral's performance. The **Bonus** 333
 334 treatment adds an additional fixed payment of \$25 to the peer who gets the referral. 334
 335 This payment is independent of the referral's actual performance (see Table 1). 335

Table 1: Incentive structure by treatment

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

¹⁰Note 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.

336 We use a between-subjects design and randomly assign half our participants to the 336
337 **Bonus** treatment. This allows us to causally identify the effect of the bonus on referral 337
338 selection. Participants learn whether their referral gets the fixed bonus before making 338
339 referral decisions. 339

340 4.4 Belief elicitation 340

341 We collect two sets of incentivized beliefs to assess the accuracy of participants' knowl- 341
342 edge about exam performance. Participants first report beliefs about their own percentile 342
343 ranking in the university for each exam area. After making referrals, participants report 343
344 their beliefs about their referrals' percentile ranking in the university. For both belief 344
345 elicitation tasks, participants earn \$5 per correct belief if their guess is within 7 per- 345
346 centiles of the true value. This margin of error is designed to balance precision with the 346
347 difficulty of the task. 347

348 5 Sample, Incentives, and Procedure 348

349 We invited all 4,417 undergraduate students who had completed their first year at the 349
350 university at the time of recruitment to participate in our experiment. A total of 837 350
351 students participated in the data collection (19% response rate). Our final sample con- 351
352 sists of 734 individuals who referred peers with whom they had taken at least one class 352
353 together (88% success rate). We randomly allocated participants to either **Baseline** or 353
354 **Bonus** treatments. 354

355 Table 2 presents key demographic characteristics and academic performance indi- 355
356 cators across treatments (see Appendix Table A.1 for selection). The sample is well- 356
357 balanced between the **Baseline** and **Bonus** conditions and we observe no statistically 357
358 significant differences in any of the reported variables (all p values > 0.1). Our sample is 358
359 characterized by a majority of middle-SES students with about one-tenth of the sample 359
360 being high-SES students. The test scores and GPA distributions are balanced. On av- 360
361 erage, participants had taken 3.8 courses together with members of their network, and 361

362 the average network consisted of 175 peers.

362

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.

363 The experiment was conducted online through Qualtrics, with participants recruited 363
364 from active students. To ensure data quality while managing costs, we randomly selected 364
365 one in ten participants for payment. Selected participants received a fixed payment of 365
366 \$17 for completion. They also received potential earnings from one randomly selected 366
367 belief question (up to \$5) and one randomly selected referral question (up to \$60). This 367
368 structure resulted in maximum total earnings of \$82. The average time to complete 368
369 the survey was 30 minutes, with an average compensation of \$80 for the one in ten 369
370 participants randomly selected for payment. Payment processing occurred through bank 370
371 transfer within 15 business days of participation. 371

372 **6 Results**

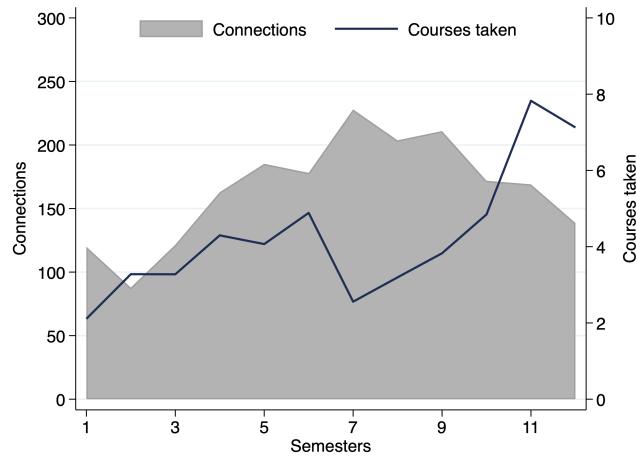
372

373 **6.1 Network characteristics**

373

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

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

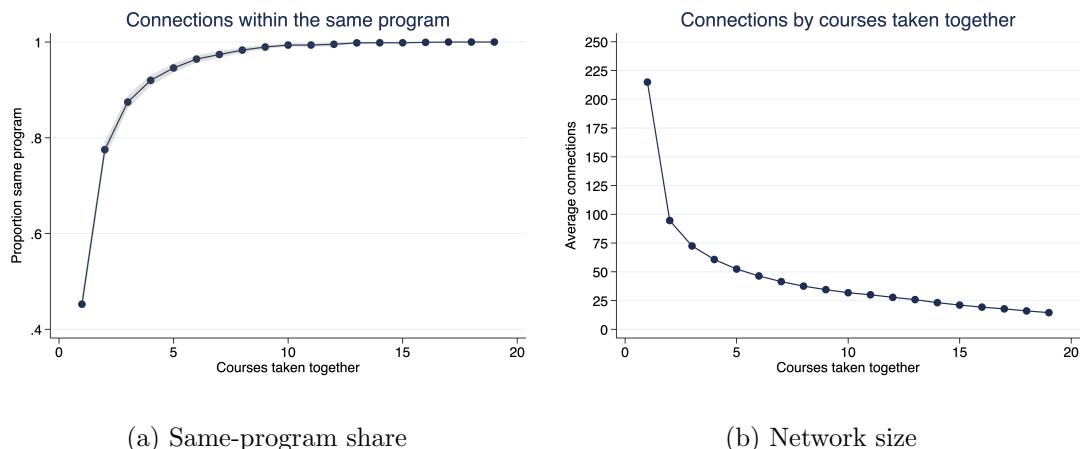


Note: This figure displays the average number of connections in blue and the average number of courses taken together with connections in grey across semesters completed. Network size (nb. of connections) peaks around 7 semesters before declining as students graduate. Connection intensity (nb. of courses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

379 We now examine how connection intensity relates to network size and composition. 379
380 First, if two students take more courses together, it is very likely that they are in the 380
381 same academic program. We plot this relationship in Figure 6a: As students take more 381
382 than 5 courses together, the share of students in their enrollment network from the same 382
383 academic program quickly exceeds 90%. Second, because students sort into specialized 383

384 academic programs, increases in courses taken together should result in decreases in 384
 385 connections. We plot this relationship in Figure 6b: As students take more than 5 385
 386 courses together, the size of their enrollment network drops dramatically from above 210 386
 387 to below 50. These patterns reveal that while participants' overall networks are large 387
 388 with relatively few courses taken together on average, they spend most of their time at 388
 389 the university within smaller, more intensive groups of peers from the same academic 389
 390 program. 390

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.

391 6.2 Referral characteristics 391

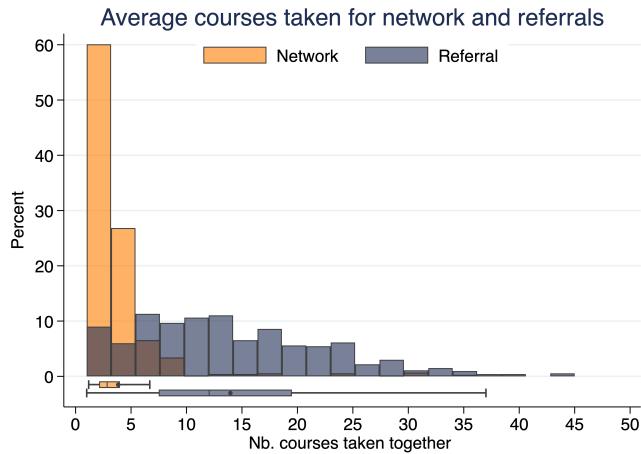
392 Participants made one referral for math and one referral for the reading part of the 392
 393 university entry exam from their enrollment networks. We observe 1,342 referrals from 393
 394 734 participants in our final dataset. More than 90% of these consist of participants 394

395 referring for both exam areas (see Appendix Table A.2). About 70% of these referrals 395
396 go to two separate individuals. We compare the outcomes across exam areas for referrals 396
397 only going to separate individuals in Appendix Table A.3 and all referrals in Appendix 397
398 Table A.4. In both cases, we find no meaningful differences between referrals made for 398
399 Math or Reading areas of the entry exam. As referrals in both exam areas come from 399
400 the same enrollment network, we pool referrals per participant and report their averages 400
401 in our main analysis to avoid inflating statistical power in our comparisons. 401

402 What are the characteristics of the individuals who receive referrals, and how do 402
403 they compare to others in the enrollment network? Because we have an entire pool of 403
404 potential candidates with one referral chosen from it, we compare the distributions for 404
405 our variables of interest between the referred and non-referred students. 405

406 First, referrals go to peers with whom the referrer has taken around 14 courses with 406
407 on average, compared to almost 4 on average with others in their network (see Figure 407
408 7). This difference of 10.1 courses is significant ($t = 34.98, p < 0.001$), indicating 408
409 that referrers choose individuals with whom they have stronger ties. While the median 409
410 referral recipient has taken 12 courses together with the referrer, the median network 410
411 member has shared only 2.8 courses. The interquartile range for referrals spans from 7.5 411
412 to 19.5 courses, compared to just 2.1 to 4.0 courses for the broader network, highlighting 412
413 the concentration of referrals among peers with higher connection intensity. In addition, 413
414 93% of referrals go to students in the same program as the referrer. 414

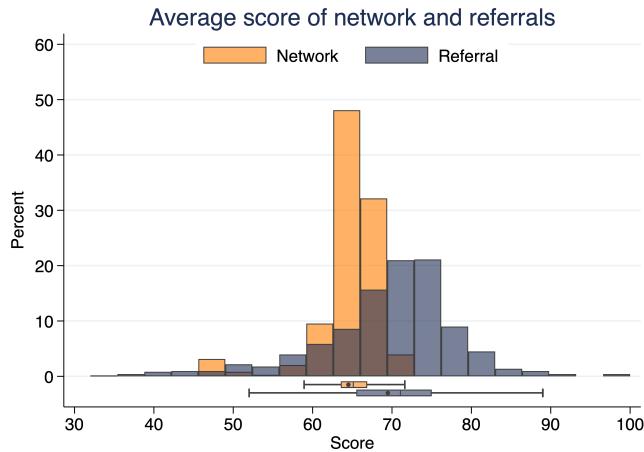
Figure 7: Courses taken together with network members and referrals



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

415 Second, we examine entry exam score differences between referred students and the 415
 416 broader network. Referrals go to peers with an average score of 69.5 points, compared 416
 417 to 64.5 points for other network members (see Figure 8). This difference of 5 points 417
 418 is significant ($t = 18.97$, $p < 0.001$), indicating that referrers choose higher-performing 418
 419 peers. While the median referral recipient scores 71 points, the median network member 419
 420 scores 65.1 points. The interquartile range for referrals spans from 65.5 to 75 points, 420
 421 compared to 63.5 to 66.9 points for the broader network, highlighting the concentration 421
 422 of referrals among higher performing peers. 422

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 $D = 71.16$, $p < 0.001$).

423 6.3 Effect of the Bonus treatment

424 Do referrals across treatments have different outcomes? We compare the performance 424
 425 and the number of courses taken together with the referrer between the **Baseline** and 425
 426 **Bonus** treatments in Table 3. We find that the the number of courses taken together 426
 427 with referrer, as well as performance measures across Reading, Math, and GPA are 427
 428 similar across treatments. Taken together, the similarities in academic performance 428
 429 and contact intensity suggest these two factors drive referrals regardless of treatment. 429
 430 For this reason, in the remainder of the paper, we report pooled results combining the 430
 431 averages of referral outcomes across treatments. 431

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
Observations	382	352	

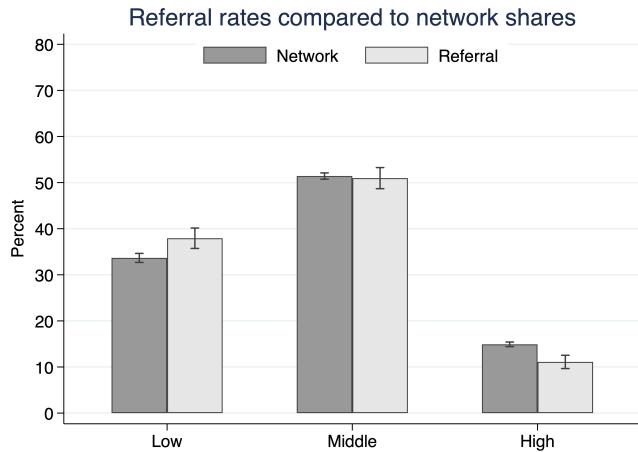
Note: This table compares the characteristics of network members who were referred under baseline vs bonus treatment conditions. *p*-values for binary outcomes are from two-sample tests of proportions; for continuous variables, from two-sample *t*-tests with unequal variances. All reported *p*-values 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. Both columns only include network members who were referred in each treatment.

432 6.4 Referral SES composition

433 To motivate the SES related biases in referral selection, we now examine the overall SES 433
 434 composition of referrals compared to the average network availability. Descriptively, 434
 435 referral patterns largely mirror underlying network structure.¹¹ Referrals to low-SES 435
 436 peers constitute 37.9% of all referrals compared to 33.7% network representation, middle- 436
 437 SES referrals account for 51.0% versus 51.4% network share, and high-SES referrals 437
 438 represent 11.1% compared to 14.9% network availability (see Figure 9). The largest 438
 439 deviation is less than 5 percentage points for any SES group. 439

¹¹Because we estimate the share of SES groups in every individual network, we get very precise estimates of the actual means. However, it is important to note that these are not independent observations. Each enrollment network is a draw with replacement from the same pool of university population, from which we calculate the proportion of SES groups per individual network, and take the average over an SES group. Pooling over SES groups who are connected with similar others systematically reduces variance (similar to resampling in bootstrapping). For this reason we choose not reporting test results in certain sections including this one and focus on describing the relationships between SES groups.

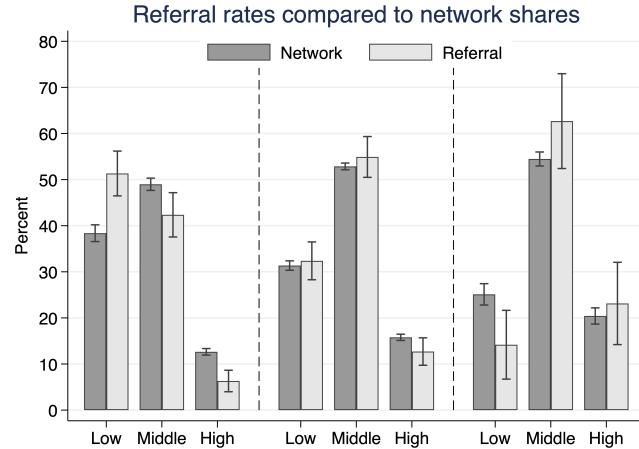
Figure 9: Referral patterns compared to network composition



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

440 Examining patterns by referrer SES reveals larger deviations. Low-SES referrers 440
441 have the largest same-SES deviation, referring 12.9 percentage points more to low-SES 441
442 students than their network composition suggests, while high-SES referrers under-refer to 442
443 low-SES students by 10.9 percentage points (see Figure 10). These descriptive findings 443
444 suggest that referral behavior diverges most from underlying network structure when 444
445 SES differences are most pronounced and motivate our formal analysis. 445

Figure 10: Referral patterns by referrer SES compared to network composition



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

446 6.5 Identifying the SES bias in referrals 446

447 We now analyze the results of the regression specification in Equation 1 and describe our 447
 448 findings in Table 4. We run three separate regressions, one for each referrer SES group, 448
 449 with a single regressor which is the referral candidate's SES. Controlling for network 449
 450 composition, we find that low-SES participants are more likely to refer other low-SES, 450
 451 and are less likely to refer high-SES relative to the probability of referring middle-SES 451
 452 peers. In contrast, we find that high-SES participants are less likely to refer other 452
 453 low-SES, relative to the probability of referring middle-SES peers. 453

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.

454 Next, we include a control for connection intensity. We proceed by adding the stan- 454
 455 dardized number of courses taken together as a control in our specification and describe 455
 456 the results in Table 5. A one standard deviation increase in the number of courses taken 456
 457 together proves to be highly significant across all models, with coefficients ranging from 457
 458 0.856 to 1.049, indicating that intensity of contact substantially increase the probability 458
 459 of referral. The high χ^2 statistics suggest that the model with this regressor provides a 459
 460 better fit than a model without. Nevertheless, low-SES participants still show a strong 460
 461 same-SES bias relative to referring middle-SES peers at the average number of courses 461
 462 taken together. This same-SES bias is not observed among middle-SES or high-SES 462
 463 referrers, who also display no statistically significant bias toward low-SES candidates. 463
 464 No referrer group shows a positive bias for high-SES candidates relative to middle-SES 464

465 candidates.

465

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

	Referrer SES		
	Low (1)	Middle (2)	High (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.

466 We then add standardized entry exam scores as a second control variable and describe 466
467 our results in Table 6. A one standard deviation increase in the entry exam score (math 467
468 and critical reading average) proves highly significant across all models, with coefficients 468
469 ranging from 0.587 to 0.883. This shows merit-based considerations due to the incentive 469
470 structure of the experiment remained central to referral decisions. The slightly higher χ^2 470
471 statistics compared to the earlier specification suggests that entry exam scores improve 471
472 model fit. The inclusion of standardized entry exam scores strengthens SES biases: Low- 472
473 SES referrers maintain their same-SES bias, with now a significant negative bias against 473

474 high-SES. Middle-SES referrers, previously showing no SES bias, now show marginal 474
475 negative bias against high-SES. Finally, high-SES referrers exhibit marginal negative 475
476 bias against low-SES candidates. 476

477 The evidence of a bias becoming significant when controlling for entry exam scores has 477
478 a nuanced interpretation. While at the university-level, low-SES typically score lower in 478
479 the entry exam, low-SES students appearing in high-SES networks are positively selected, 479
480 scoring about 0.14 standard deviations higher than middle-SES students (see Appendix 480
481 Table A.5). Controlling for performance thus removes this positive selection and reveals 481
482 the SES bias that was previously underestimated by above average performance of low- 482
483 SES. Vice versa, high-SES in low-SES networks perform 0.12 standard deviations better 483
484 than middle-SES students. The bias was underestimated as high-SES candidates' better 484
485 performance relative to middle-SES increased referrals. Controlling for exam scores 485
486 reveal that both high- and low-SES referrers have negative SES bias towards one another 486
487 that operates independently of – and counter to – performance-based considerations. 487
488 What makes a symmetric bias interpretation difficult is that while biased against low- 488
489 SES, high-SES referrers do not (under any specification) display a positive bias towards 489
490 their in-group. 490

491 We conclude that the SES bias in referral selection is primarily driven by low-SES 491
492 referrers who exhibit strong in-group preferences. Middle- and high-SES referrers show 492
493 no systematic discrimination against other SES groups once we account for network 493
494 composition and other relevant factors contributing to the referral decision. We will 494
495 explore potential mechanisms that help explain this unexpected result. 495

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.

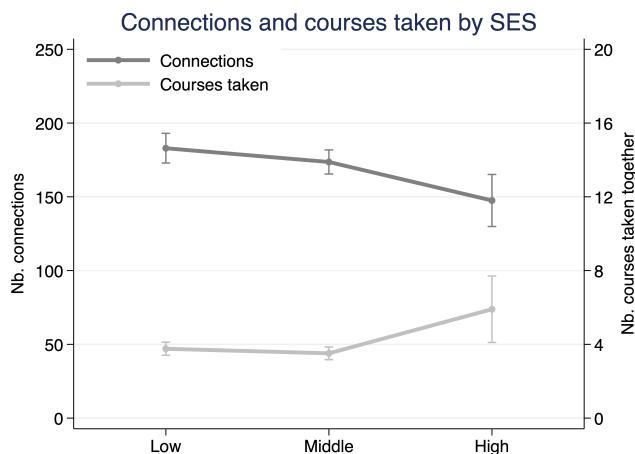
496 7 Potential Mechanisms and Robustness Checks 496

497 7.1 SES diversity in networks 497

498 How do enrollment networks differ across SES groups? We look at how the number 498
 499 of connections (network size) and number of courses taken together (contact intensity) 499

500 change across SES groups in Figure 11. Low- and middle-SES students have larger 500
 501 networks but take fewer courses together with network members, while high-SES stu- 501
 502 dents have smaller, denser networks. Specifically, both low- and middle-SES students 502
 503 have significantly larger networks than high-SES students ($t = 3.03, p = 0.003$ and 503
 504 $t = 2.49, p = 0.013$, respectively), but high-SES students take significantly more courses 504
 505 with their network members than both low- ($t = -3.70, p < .001$) and middle-SES 505
 506 ($t = -4.20, p < .001$). 506

Figure 11: Network size and courses taken together by SES



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

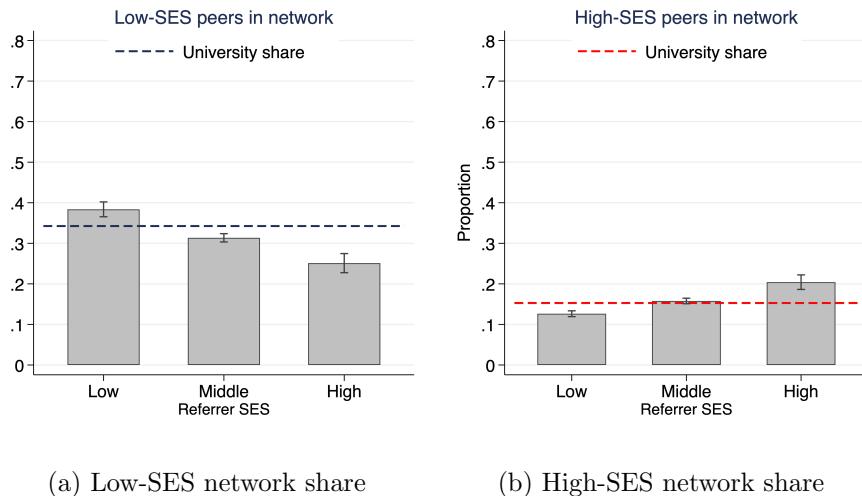
507 What are the diversity-related consequences of SES-driven differences across net- 507
 508 works? In terms of network compositions, participants could connect with other SES 508
 509 groups at different rates than would occur randomly depending on their own SES. We il- 509
 510 lustrate the average network shares conditional on referrer SES for low-SES in Figure 12a 510
 511 and for high-SES in Figure 12b.¹² We observe modest deviations from university-wide 511

¹²For sake of brevity we omit middle-SES from this exposition. For the complete relationship, see Appendix Figure A.3.

512 SES shares in enrollment networks: Low-SES referrers have on average 38.4% low-SES 512
 513 peers compared to the university average of 34.3%, while high-SES referrers have 20.4% 513
 514 high-SES connections compared to the university average of 15.3%. 514

515 We find larger differences when studying connections between SES groups: Low- 515
 516 SES referrers connect with other low-SES at much higher rates than high-SES referrers 516
 517 (38.4% vs 25.1%). Conversely, high-SES referrers connect more with other high-SES 517
 518 than low-SES referrers (20.4% vs 12.6%). Middle-SES referrers are in between the two 518
 519 extreme patterns, connecting with middle-SES at higher rates than low-SES referrers 519
 520 (52.9% vs 49.0%) but lower rates than high-SES referrers (52.9% vs 54.5%). These 520
 521 findings indicate SES-based segregation in networks, with same-SES homophily across 521
 522 groups. 522

Figure 12: Network shares of SES groups



(a) Low-SES network share

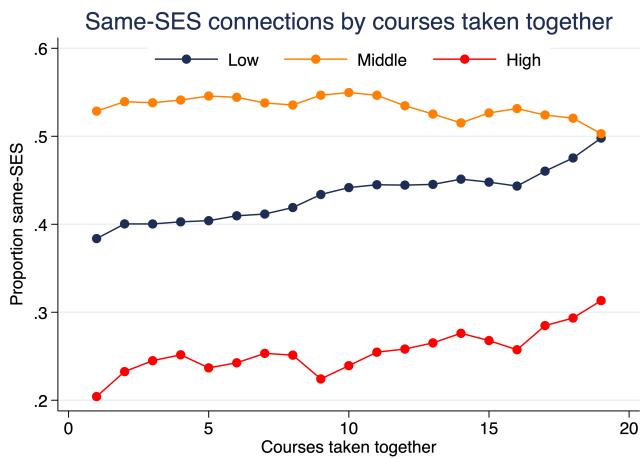
(b) High-SES network share

Note: Figures illustrate the average network shares of low- and high- SES peers conditional on referrer SES. Horizontal lines plot the university-wide shares of SES groups (Low: 35%, High: 14%). While the share of low-SES peers in the network decreases as referrer SES increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

523 While same-SES students are connected more often with each other, so far we did 523

524 not look at the consequences in terms of number of courses taken together with network 524
 525 members. What are the diversity implications of increased connection intensity between 525
 526 students? As students take more courses together with peers, the share of same-SES 526
 527 peers in the networks of low- and high-SES increases while the share of middle-SES 527
 528 declines (see Figure 13). Both increases are substantial, amounting to 50% for high-, and 528
 529 30% for low-SES. Considering that beyond 5 courses taken together network members 529
 530 are almost entirely within the same program, these suggest program selection may have 530
 531 strong consequences for SES diversity in our setting. 531

Figure 13: Network size and connection intensity



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.

532 7.2 SES diversity in referral choice sets 532

533 How did the referrer choice sets look like in practice? We combine our findings from 533
 534 network diversity and its relationship with connection intensity, together with referral 534
 535 selection. In Section 6.2, we found that referrals went to peers with whom the median 535
 536 participant took 12 courses (average 14). By restricting the networks for courses taken 536

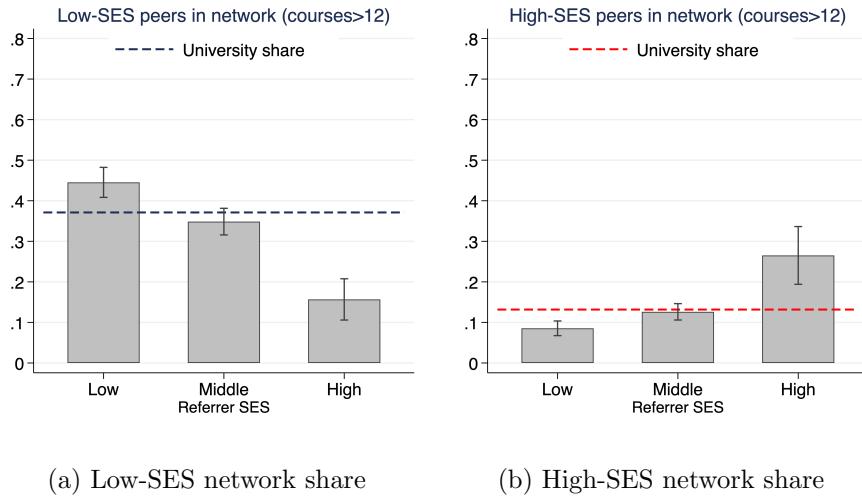
537 above the median, we get an *ex post* snapshot of referrer choice sets. 537

538 We show the average network shares conditional on referrer SES and above median 538
539 number of courses taken together for low-SES in Figure 14a and for high-SES in Figure 539
540 14b.¹³ Network compositions above the median number of courses taken reveal strong 540
541 segregation effects in referral choice sets: Low-SES networks contain 44.5% low-SES 541
542 peers, higher than the 35% university-wide share by 9.5 percentage points. Conversely, 542
543 high-SES students are under-represented in low-SES networks at only 8.6% average 543
544 share, compared to the 14% population share (−5.4 pp.). At the other extreme, high-SES 544
545 networks show the reverse pattern with average low-SES share dropping to just 15.7%, 545
546 a 19.3 percentage point decrease relative to the university average. High-SES students 546
547 have a same-SES concentration at 26.5%, doubling their 14% population share (+12.5 547
548 pp.). Middle-SES networks remain relatively balanced and closely track population 548
549 proportions. 549

550 Put differently, in an environment where 1 out of 3 students are low-SES, the chance 550
551 that low-SES are considered for a referral by high-SES at random is already less than 551
552 1/6. This confirms that low-SES and high-SES practically have non-overlapping net- 552
553 works despite having opportunities to meet on equal status students at the university. 553
554 While referral selection being driven by availability and performance is positive, network 554
555 segregation has such a large impact on diversity. We now explore program selection as 555
556 a key driver. 556

¹³In Appendix Figure A.4 we present the complete relationship including middle-SES.

Figure 14: Network shares of SES groups above median connection intensity

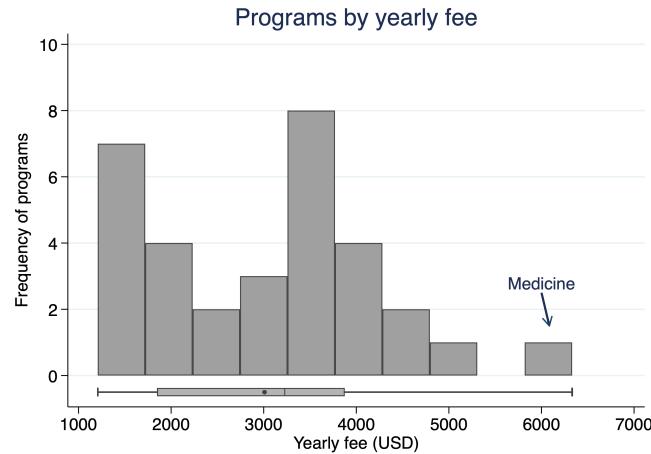


Note: Figures illustrate the average network shares of low- and high- SES peers conditional on referrer SES above the median number of courses taken together. Horizontal lines plot the university-wide shares of SES groups (Low: 35%, High: 14%). While the share of low-SES peers in the network decreases as referrer SES increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

557 7.3 Program selection and SES diversity

558 Academic programs at this university use cost-based pricing, and typically less than 5% 558
 559 of students receive any kind of scholarship. Based on this, we first calculate how much 559
 560 every undergraduate program at the university is expected to cost students per year 560
 561 (see Figure 15). Considering that net minimum monthly wage stands at \$200 and the 561
 562 average Colombian salary around \$350, the cost differences between programs are large 562
 563 enough to make an impact on program selection. Is it the case that SES groups select 563
 564 into programs with financial considerations? 564

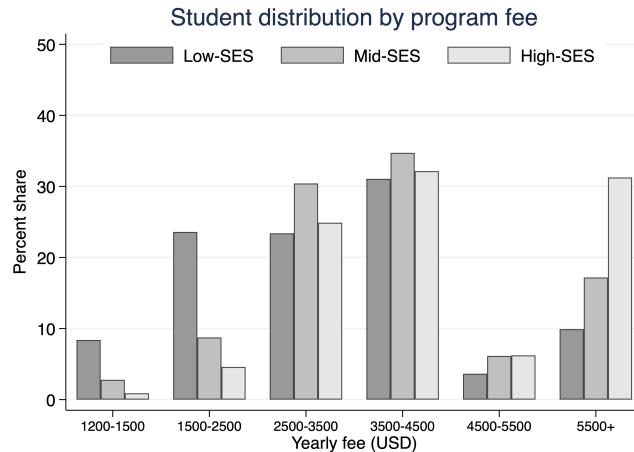
Figure 15: Undergraduate programs sorted by fee



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

We then examine how SES groups are distributed across programs to identify evidence of SES-based selection (see Figure 16). Indeed, low-SES students select into more affordable programs, followed by middle-SES students. High-SES students sort almost exclusively into above-average costing programs, with a third selecting into medicine and creating a very skewed distribution. The distributions are significantly different across all pairwise comparisons: low-SES vs. middle-SES ($D = 33.89, p < 0.001$), 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 16: SES distribution by program fee



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

577 7.4 Robustness check: Contact intensity and sharing academic pro- 578 grams

579 Does the number of courses taken together have an independent effect that goes beyond 579
580 identifying peers in the same academic program? To evaluate this question we leverage 580
581 our administrative data, and identify peers within the same program: In each individ- 581
582 ual network we observe the participant-specific academic program for the participant 582
583 making the referral and alternative-specific academic program for each referral candi- 583
584 date. We add this new variable in our specification and describe our findings in Table 584
585 7. Being in the same academic program has a substantial positive effect on referral 585
586 likelihood, with coefficients ranging from 1.257 to 2.198 across all referrer SES groups. 586
587 This confirms that program affiliation serves as a strong predictor of referral decisions, 587
588 reflecting increased familiarity. Our comparison of interest is the point estimate for the 588
589 standardized number of courses taken. Across all three referrer groups, the standardized 589
590 number of courses taken together maintains its statistical significance after controlling 590

591 for same program membership. The coefficient magnitudes are expectedly smaller com- 591
592 pared to specifications without program controls (ranging from 0.688 to 0.930) as the 592
593 newly added variable is a moderator: Matching academic programs leads to taking more 593
594 courses together. The remaining estimates in our model remain robust to the inclusion 594
595 of the same-program variable with little change in point estimates. The persistence of 595
596 statistical significance (all $p < 0.001$) suggests that the number of courses taken together 596
597 has an independent effect on referral decisions. To sum, our measure of contact inten- 597
598 sity seems to capture meaningful social interaction patterns that lead to referrals, and 598
599 go beyond simply identifying matching academic programs. 599

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.

600 **8 Conclusion**

600

601 In this paper, we study whether SES groups are biased toward one another beyond 601
602 what is attributable to differences in their networks, and the effects of different incentive 602
603 structures on referral behavior. Through a lab-in-the-field experiment that leverages 603
604 enrollment networks at a socially diverse university, we find that the SES biases in 604
605 referrals originate mostly from network structures, and referrals under performance-pay 605
606 incentives do not exacerbate existing SES inequalities. 606

607 Our findings reveal that enrollment networks are surprisingly segregated and referrals 607
608 from these networks reflect closely the choice sets of the referrers. We identify program 608
609 selection as the key mechanism driving this segregation. Low-SES students select into 609
610 more affordable programs, and program selection plays a major part in segregating 610
611 SES groups where low- and high-SES take more courses with their own SES group. 611
612 Consequently, referrals come almost exclusively from the same academic program as the 612
613 referrer, limiting the SES-diversity of their choice sets. Regardless of the bonus for the 613
614 referral candidate, participants also pick higher performing peers with whom they have 614
615 taken many courses together. We find that only low-SES referrers exhibit a same-SES 615
616 bias. These findings suggest that the underlying network structure plays a crucial role 616
617 in referrals, where institutional action can remedy the network segregation. 617

618 These results complement the broader literature where much of the bias in referrals 618
619 can be attributable to the “practical” choice sets of the referrers. While previous work 619
620 demonstrates that about half of referrals come from a smaller, elicited network of close 620
621 friends ([Hederos et al., 2025](#)), we go the other way and use administrative data to 621
622 construct a complete network which presumably includes close social relationships at the 622
623 institutional level. Having access to the complete network thus eliminates any potential 623
624 for under or overestimating taste-based biases ([Griffith, 2022](#)). Under performance-pay 624
625 incentives, referrers identify productive others regardless of additional financial rewards 625
626 for the referral candidate. Still, the lack of a treatment effect suggests that in both 626
627 incentive structures referrers pick close ties, shifting the responsibility to institutional 627

628 actors to create diverse environments where cross-SES social interaction can take place 628
629 more frequently and allow more diversity in networks. 629

630 These findings have policy implications. Looking forward, institutions can play a 630
631 crucial role in achieving SES equality of opportunity in higher education. Universities 631
632 are already a setting in which low-SES get exposed to typically a higher than population 632
633 share of higher-SES individuals than at other settings (Chetty et al., 2022b). Yet, 633
634 segregation within the higher education institutions remain a source for SES inequality. 634
635 If low-SES peers never get to interact in meaningful ways with higher-SES, e.g., by 635
636 taking courses together, the premise of social mobility through social channels remains 636
637 severely underexploited. Future studies should work on ways to reduce SES segregation 637
638 in collaboration with institutions, where having access to complete enrollment networks 638
639 in addition to the typical friendship elicitation methods could help identifying the exact 639
640 overlap between the two distinct approaches. 640

641 References

641

- 642 Alan, S., Duysak, E., Kibilay, E., & Mumcu, I. (2023). Social Exclusion and Ethnic 642
643 Segregation in Schools: The Role of Teachers' Ethnic Prejudice. *The Review of 643
644 Economics and Statistics*, 105(5), 1039–1054. doi: 10.1162/rest_a_01111 644
- 645 Angulo, R., Gaviria, A., Páez, G. N., & Azevedo, J. P. (2012). Movilidad social en 645
646 colombia. *Documentos CEDE*. 646
- 647 Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the 647
648 workplace: Evidence from personnel data. *Econometrica*, 77(4), 1047–1094. 648
- 649 Beaman, L., Keleher, N., & Magruder, J. (2018). Do Job Networks Disadvantage 649
650 Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor 650
Economics*, 36(1), 121–157. doi: 10.1086/693869 651
- 652 Beaman, L., & Magruder, J. (2012). Who Gets the Job Referral? Evidence from a 652
653 Social Networks Experiment. *American Economic Review*, 102(7), 3574–3593. 653
654 doi: 10.1257/aer.102.7.3574 654
- 655 Bolte, L., Immorlica, N., & Jackson, M. O. (2021). *The Role of Referrals in Immobility, 655
Inequality, and Inefficiency in Labor Markets*. (Working Paper) 656
- 657 Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory 657
and research for the sociology of education* (pp. 241–258). New York: Greenwood 658
659 Press. 659
- 660 Brown, M., Setren, E., & Topa, G. (2016). Do informal referrals lead to better matches? 660
661 evidence from a firm's employee referral system. *Journal of Labor Economics*, 661
662 34(1), 161–209. 662
- 663 Calvo-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on em- 663
664 ployment and inequality. *American economic review*, 94(3), 426–454. 664
- 665 Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., 665
666 ... Wernerfelt, N. (2022a). Social capital 1: Measurement and associations with 666
667 economic mobility. *Nature*, 608(7921), 108–121. doi: 10.1038/s41586-022-04996-4 667
- 668 Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., ... 668

- 669 Wernerfelt, N. (2022b). Social capital 2: Determinants of economic connectedness. 669
670 *Nature*, 608(7921), 122–134. doi: 10.1038/s41586-022-04997-3 670
- 671 Díaz, J., Munoz, M., Reuben, E., & Tuncer, R. (2025, March). *Peer skill identification* 671
672 *and social class: Evidence from a referral field experiment.* (Working Paper) 672
- 673 Dustmann, C., Glitz, A., Schönberg, U., & Brücker, H. (2016). Referral-based job search 673
674 networks. *The Review of Economic Studies*, 83(2), 514–546. 674
- 675 Fergusson, L., & Flórez, S. A. (2021a). Desigualdad educativa en colombia. In 675
676 J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De* 676
677 *cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* 677
678 Bogotá: Ariel. 678
- 679 Fergusson, L., & Flórez, S. A. (2021b). Distinción escolar. In J. C. Cárdenas, L. Fer- 679
680 gusson, & M. García Villegas (Eds.), *La quinta puerta: De cómo la educación en* 680
681 *colombia agudiza las desigualdades en lugar de remediarlas.* Bogotá: Ariel. 681
- 682 Friebel, G., Heinz, M., Hoffman, M., & Zubanov, N. (2023). What do employee referral 682
683 programs do? measuring the direct and overall effects of a management practice. 683
684 *Journal of Political Economy*, 131(3), 633–686. 684
- 685 García, S., Rodríguez, C., Sánchez, F., & Bedoya, J. G. (2015). La lotería de la 685
686 cuna: La movilidad social a través de la educación en los municipios de colombia. 686
687 *Documentos CEDE*. 687
- 688 García Villegas, M., & Cobo, P. (2021). La dimensión cultural del apartheid educativo. 688
689 In J. C. Cárdenas, L. Fergusson, & M. García Villegas (Eds.), *La quinta puerta: De* 689
690 *cómo la educación en colombia agudiza las desigualdades en lugar de remediarlas.* 690
691 Bogotá: Ariel. 691
- 692 Gee, L. K., Jones, J. J., & Burke, M. (2017). Social networks and labor markets: How 692
693 strong ties relate to job finding on facebook's social network. *Journal of Labor* 693
694 *Economics*, 35(2), 485–518. 694
- 695 Griffith, A. (2022). Name Your Friends, but Only Five? The Importance of Censoring in 695
696 Peer Effects Estimates Using Social Network Data. *Journal of Labor Economics*. 696
697 doi: 10.1086/717935 697

- 698 Guevara S, J. D., & Shields, R. (2019). Spatializing stratification: Bogotá. *Ardeth. A* 698
699 *Magazine on the Power of the Project*(4), 223–236. 699
- 700 Hederos, K., Sandberg, A., Kvissberg, L., & Polano, E. (2025). Gender homophily 700
701 in job referrals: Evidence from a field study among university students. *Labour* 701
702 *Economics*, 92, 102662. 702
- 703 Hudson, R. A., & Library of Congress (Eds.). (2010). *Colombia: a country study* 703
704 (5th ed.). Washington, D.C: Federal Research Division, Library of Congress: For 704
705 sale by the Supt. of Docs., U.S. G.P.O. Retrieved from the Library of Congress, 705
706 <https://www.loc.gov/item/2010009203/>. 706
- 707 Jaramillo-Echeverri, J., & Álvarez, A. (2023). *The Persistence of Segregation in Edu* 707
708 *cation: Evidence from Historical Elites and Ethnic Surnames in Colombia* (SSRN 708
709 Scholarly Paper No. 4575894). Rochester, NY: Social Science Research Network. 709
710 doi: 10.2139/ssrn.4575894 710
- 711 Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. 711
712 *American Journal of Sociology*, 115(2), 405–450. Retrieved from <https://www> 712
713 [.journals.uchicago.edu/doi/abs/10.1086/599247](https://journals.uchicago.edu/doi/abs/10.1086/599247) doi: 10.1086/599247 713
- 714 Kramarz, F., & Nordström Skans, O. (2014). When strong ties are strong: Networks and 714
715 youth labour market entry. *The Review of Economic Studies*, 81(3), 1164–1200. 715
- 716 Lin, N., Ensel, W. M., & Vaughn, J. C. (1981). Social Resources and Strength of 716
717 Ties: Structural Factors in Occupational Status Attainment. *American Sociological 717
718 Review*, 46(4), 393–405. doi: 10.2307/2095260 718
- 719 Loury, G. C. (1977). A dynamic theory of racial income differences. In P. A. Wallace 719
720 & A. M. LaMond (Eds.), *Women, minorities, and employment discrimination* 720
721 (pp. 153–186). Lexington, MA: Lexington Books. (Originally published as Dis- 721
722 cussion Paper 225, Northwestern University, Center for Mathematical Studies in 722
723 Economics and Management Science, 1976) 723
- 724 McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily 724
725 in social networks. *Annual review of sociology*, 27(1), 415–444. 725
- 726 Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an 726

- 727 Economic Analysis. *American Economic Review*. 727

728 Mouw, T. (2003). Social Capital and Finding a Job: Do Contacts Matter? *American 728
Sociological Review*, 68(6), 868–898. doi: 10.1177/000312240306800604 729

729 Pallais, A., & Sands, E. G. (2016). Why the Referential Treatment? Evidence from 730
Field Experiments on Referrals. *Journal of Political Economy*, 124(6), 1793–1828. 731

730 doi: 10.1086/688850 732

732 Pedulla, D. S., & Pager, D. (2019). Race and networks in the job search process. 733
American Sociological Review, 84, 983–1012. doi: 10.1177/0003122419883255 734

734 Rohrer, J. M., Keller, T., & Elwert, F. (2021). Proximity can induce diverse friendships: 735
A large randomized classroom experiment. *PLOS ONE*, 16(8), e0255097. doi: 736

736 10.1371/journal.pone.0255097 737

737 Smith, S. S. (2005). “Don’t put my name on it”: Social Capital Activation and Job- 738
Finding Assistance among the Black Urban Poor. *American Journal of Sociology*, 739
111(1), 1–57. doi: 10.1086/428814 740

740 Stansbury, A., & Rodriguez, K. (2024). The class gap in career progression: Evidence 741
from US academia. *Working Paper*. 742

742 Topa, G. (2019). Social and spatial networks in labour markets. *Oxford Review of 743
Economic Policy*, 35(4), 722–745. 744

744 United Nations. (2023). *Social panorama of latin america and the caribbean 745
2023: labour inclusion as a key axis of inclusive social development*. 746
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> 747
748 749

750 Uribe-Mallarino, C. (2008). Estratificación social en bogotá: de la política pública a la 750
dinámica de la segregación social. *Universitas humanistica*(65), 139–172. 751

751 Wang, S.-Y. (2013). Marriage networks, nepotism, and labor market outcomes in china. 752
American Economic Journal: Applied Economics, 5(3), 91–112. 753

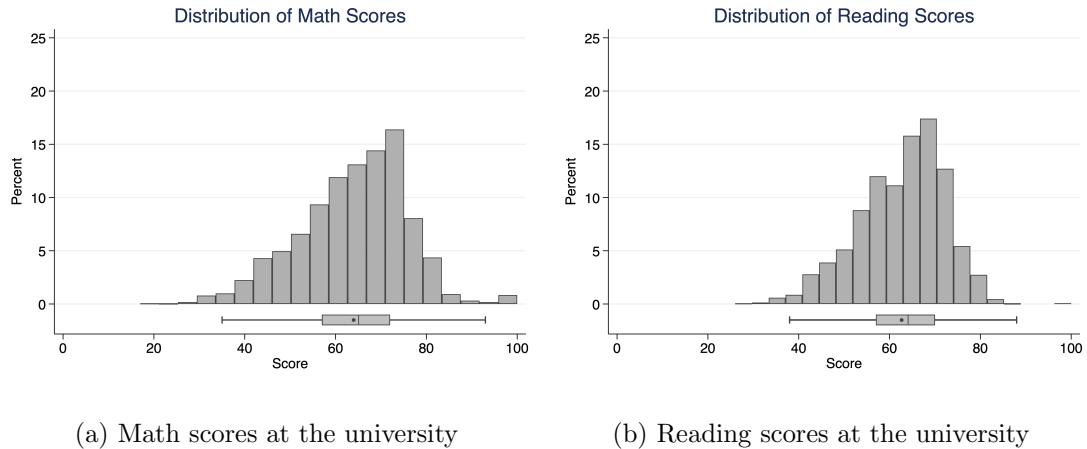
753 Witte, M. (2021). Why do workers make job referrals? experimental evidence from 754
ethiopia. *Working Paper*. 755

756 World Bank. (2024). *Regional poverty and inequality update spring 2024* 756
757 (Poverty and Equity Global Practice Brief). Washington, D.C.: World 757
758 Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/099070124163525013/P17951815642cf06e1aec4155e4d8868269> 758
759 759

761 Additional Figures

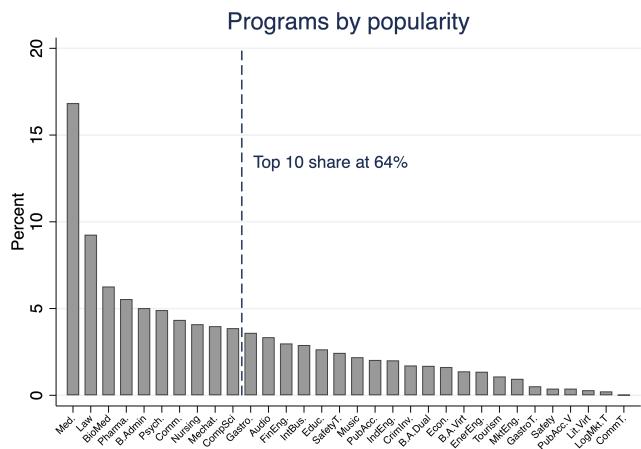
761

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



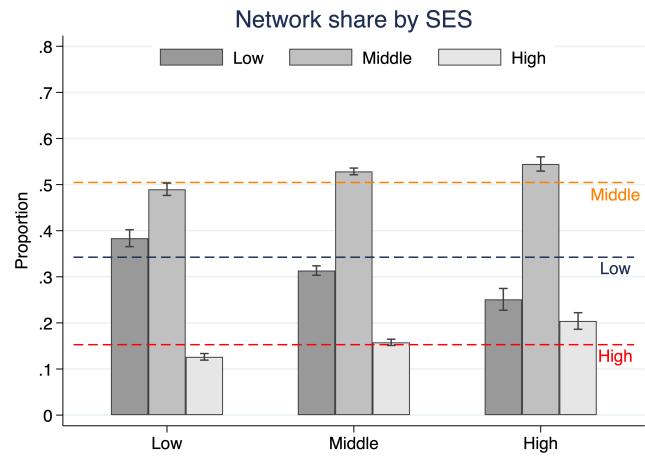
Note: Reading scores (left panel) and math scores (right panel) show tight distributions with approximately 75% of students falling within just 13-15 points of each other.

Figure A.2: Distribution of students across undergraduate programs



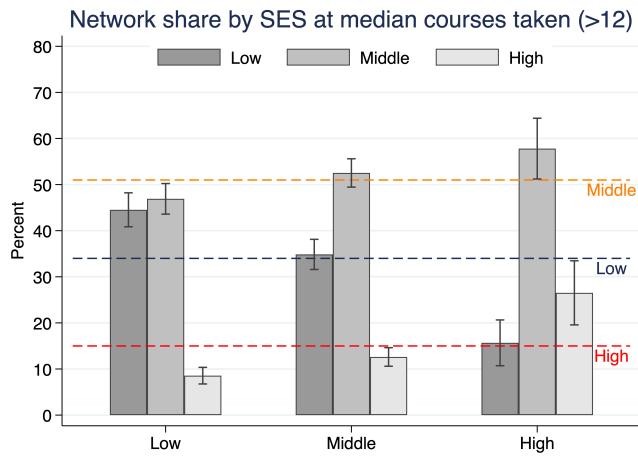
Note: This figure shows the concentration of students across 32 undergraduate programs at the university. Students cluster around certain programs. The top 5 most popular programs (Medicine, Law, Biomedical Engineering, Pharmacy Technology, and Business Administration) account for 43% of all undergraduates, and the top 10 most popular programs account for 63% of students.

Figure A.3: Network shares by SES



Note: This figure displays the average network shares of SES groups respectively for low-, middle-, and high-SES referrers. Horizontal lines plot the university-wide shares of each SES group (Low: 35%, Mid: 51%, High: 14%). While the share of low-SES peers in the network decreases as the SES of the referrers increases, the share of high-SES peers in the network increases. Error bars represent 95% confidence intervals.

Figure A.4: Network shares by SES at courses taken above 12



Note: This figure displays the average network shares of SES groups respectively for low-, middle-, and high-SES referrers above the median number of courses taken together. Horizontal lines plot the university-wide shares of each SES group (Low: 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.

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

763 **B Experiment**

763

764 We include the English version of the instructions used in Qualtrics. Participants saw 764
765 the Spanish version. Horizontal lines in the text indicate page breaks and clarifying 765
766 comments are inside brackets. 766

767 **Consent**

767

768 You have been invited to participate in this decision-making study. This study is directed 768
769 by [omitted for anonymous review] and organized with the support of the Social Bee Lab 769
770 (Social Behavior and Experimental Economics Laboratory) at UNAB. 770

771 In this study, we will pay **one (1)** out of every **ten (10)** participants, who will be 771
772 randomly selected. Each selected person will receive a fixed payment of **70,000** (seventy 772
773 thousand pesos) for completing the study. Additionally, they can earn up to **270,000** 773
774 (two hundred and seventy thousand pesos), depending on their decisions. So, in total, 774
775 if you are selected to receive payment, you can earn up to **340,000** (three hundred and 775
776 forty thousand pesos) for completing this study. 776

777 If you are selected, you can claim your payment at any Banco de Bogotá office by 777
778 presenting your ID. Your participation in this study is voluntary and you can leave the 778
779 study at any time. If you withdraw before completing the study, you will not receive 779
780 any payment. 780

781 The estimated duration of this study is 20 minutes. 781

782 The purpose of this study is to understand how people make decisions. For this, we will 782
783 use administrative information from the university such as the SABER 11 test scores of 783
784 various students (including you). Your responses will not be shared with anyone and your 784
785 participation will not affect your academic records. To maintain strict confidentiality, the 785
786 research results will not be associated at any time with information that could personally 786

787 identify you.

787

788 There are no risks associated with your participation in this study beyond everyday risks. 788
789 However, if you wish to report any problems, you can contact Professor [omitted for 789
790 anonymous review]. For questions related to your rights as a research study participant, 790
791 you can contact the IRB office of [omitted for anonymous review]. 791

792 By selecting the option “I want to participate in the study” below, you give your con- 792
793 sent to participate in this study and allow us to compare your responses with some 793
794 administrative records from the university. 794

795 • I want to participate in the study [advances to next page] 795

796 • I do not want to participate in the study 796

797 _____ 797

798 Student Information 798

799 Please write your student code. In case you are enrolled in more than one program 799
800 simultaneously, write the code of the first program you entered: 800

801 [Student ID code] 801

802 What semester are you currently in? 802

803 [Slider ranging from 1 to 11] 803

804 _____ 804

805 [Random assignment to treatment or control] 805

806 **Instructions**

806

807 The instructions for this study are presented in the following video. Please watch it 807
808 carefully. We will explain your participation and how earnings are determined if you are 808
809 selected to receive payment. 809

810 [Treatment-specific instructions in video format] 810

811 If you want to read the text of the instructions narrated in the video, press the “Read 811
812 instruction text” button. Also know that in each question, there will be a button with 812
813 information that will remind you if that question has earnings and how it is calculated, 813
814 in case you have any doubts. 814

815 • I want to read the instructions text [text version below] 815

816 —————— 816

817 In this study, you will respond to three types of questions. First, are the belief questions. 817
818 For belief questions, we will use as reference the results of the SABER 11 test that you 818
819 and other students took to enter the university, focused on three areas of the exam: 819
820 mathematics, reading, and English. 820

821 For each area, we will take the scores of all university students and order them from 821
822 lowest to highest. We will then group them into 100 percentiles. The percentile is a 822
823 position measure that indicates the percentage of students with an exam score that is 823
824 above or below a value. 824

825 For example, if your score in mathematics is in the 20th percentile, it means that 20 825
826 percent of university students have a score lower than yours and the remaining 80 percent 826
827 have a higher score. A sample belief question is: “compared to university students, in 827
828 what percentile is your score for mathematics?” 828

829 If your answer is correct, you can earn 20 thousand pesos. We say your answer is correct 829

830 if the difference between the percentile you suggest and the actual percentile of your 830
831 score is not greater than 7 units. For example, if you have a score that is in the 33rd 831
832 percentile and you say it is in the 38th, the answer is correct because the difference is 832
833 less than 7. But if you answer that it is in the 41st, the difference is greater than 7 and 833
834 the answer is incorrect. 834

835 The second type of questions are recommendation questions and are also based on the 835
836 mathematics, reading, and English areas of the SABER 11 test. We will ask you to think 836
837 about the students with whom you have taken or are taking classes, to recommend from 837
838 among them the person you consider best at solving problems similar to those on the 838
839 SABER 11 test. 839

840 When you start typing the name of your recommended person, the computer will show 840
841 suggestions with the full name, program, and university entry year of different students. 841
842 Choose the person you want to recommend. If the name doesn't appear, check that you 842
843 are writing it correctly. Do not use accents and use 'n' instead of 'ñ'. If it still doesn't 843
844 appear, it may be because that person is not enrolled this semester or because they did 844
845 not take the SABER 11 test. In that case, recommend someone else. 845

846 You can earn up to 250,000 pesos for your recommendation. We will multiply your 846
847 recommended person's score by 100 pesos if they are in the first 50 percentiles. We will 847
848 multiply it by 500 pesos if your recommended person's score is between the 51st and 848
849 65th percentile. If it is between the 66th and 80th percentile, we will multiply your 849
850 recommended person's score by 1000 pesos. If the score is between the 81st and 90th 850
851 percentile, you earn 1500 pesos multiplied by your recommended person's score. And if 851
852 the score is between the 91st and 100th percentile, we will multiply your recommended 852
853 person's score by 2500 pesos to determine the earnings. 853

854 The third type of questions are information questions and focus on aspects of your 854
855 personal life or your relationship with the people you have recommended. 855

856 **Earnings** 856

857 Now we will explain who gets paid for participating and how the earnings for this study 857
858 are assigned. The computer will randomly select one out of every 10 participants to pay 858
859 for their responses. For selected individuals, the computer will randomly choose one of 859
860 the three areas, and from that chosen area, it will pay for one of the belief questions. 860

861 Similarly, the computer will randomly select one of the three areas to pay for one of the 861
862 recommendation questions. 862

863 **Additionally, if you are selected to receive payment, your recommended per- 863
864 son in the chosen area will receive a fixed payment of 100 thousand pesos. 864
865 [Only seen if assigned to the treatment]** 865

866 Each person selected to receive payment for this study can earn: up to 20 thousand pesos 866
867 for one of the belief questions, up to 250 thousand pesos for one of the recommendation 867
868 questions, and a fixed payment of 70 thousand pesos for completing the study. 868

869 Selected individuals can earn up to 340 thousand pesos. 869

870 870

871 [Participants go through all three Subject Areas in randomized order] 871

872 **Subject Areas** 872

873 **Critical Reading** 873

874 For this section, we will use as reference the Critical Reading test from SABER 11, which 874
875 evaluates the necessary competencies to understand, interpret, and evaluate texts that 875
876 can be found in everyday life and in non-specialized academic fields. 876

877 [Clicking shows the example question from SABER 11 below] 877

878 Although the democratic political tradition dates back to ancient Greece, political 878
879 thinkers did not address the democratic cause until the 19th century. Until then, democ- 879
880 racy had been rejected as the government of the ignorant and unenlightened masses. 880
881 Today it seems that we have all become democrats without having solid arguments in 881
882 favor. Liberals, conservatives, socialists, communists, anarchists, and even fascists have 882
883 rushed to proclaim the virtues of democracy and to show their democratic credentials 883
884 (Andrew Heywood). According to the text, which political positions identify themselves 884
885 as democratic? 885

- 886 • Only political positions that are not extremist 886
887 • The most recent political positions historically 887
888 • The majority of existing political positions 888
889 • The totality of possible political currents 889

890 —————— 890

891 **Mathematics** 891

892 This section references the Mathematics test from SABER 11, which evaluates people's 892
893 competencies to face situations that can be resolved using certain mathematical tools. 893

894 [Clicking shows the example question from SABER 11 below] 894

895 A person living in Colombia has investments in dollars in the United States and knows 895
896 that the exchange rate of the dollar against the Colombian peso will remain constant 896
897 this month, with 1 dollar equivalent to 2,000 Colombian pesos. Their investment, in 897
898 dollars, will yield profits of 3% in the same period. A friend assures them that their 898
899 profits in pesos will also be 3%. Their friend's statement is: 899

- 900 • Correct. The proportion in which the investment increases in dollars is the same 900
901 as in pesos. 901

- | | | |
|-----|---|-----|
| 902 | • Incorrect. The exact value of the investment should be known. | 902 |
| 903 | • Correct. 3% is a fixed proportion in either currency. | 903 |
| 904 | • Incorrect. 3% is a larger increase in Colombian pesos. | 904 |

905 905

906 English 906

907 This section uses the English test from SABER 11 as a reference, which evaluates that 907
908 the person demonstrates their communicative abilities in reading and language use in 908
909 this language. 909

910 [Clicking shows the example question from SABER 11 below] 910

911 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.

- 916 • Conversation 2: It rained a lot last night! 916

917 – Did you accept? 917

918 – Did you understand? 918

919 – Did you sleep? 919

[Following parts are identical for all Subject Areas and are not repeated here for brevity] 921

922 **Your Score**

922

923 Compared to university students, in which percentile do you think your [Subject Area] 923
924 test score falls (1 is the lowest percentile and 100 the highest)? 924

925 [Clicking shows the explanations below] 925

926 How is a percentile calculated? 926

927 A percentile is a position measurement. To calculate it, we take the test scores for all 927
928 students currently enrolled in the university and order them from lowest to highest. The 928
929 percentile value you choose refers to the percentage of students whose score is below 929
930 yours. For example, if you choose the 20th percentile, you're indicating that 20% of 930
931 students have a score lower than yours and the remaining 80% have a score higher than 931
932 yours. 932

933 What can I earn for this question? 933

934 For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the 934
935 difference between your response and the correct percentile is less than 7. For example, if 935
936 the percentile where your score falls is 33 and you respond with 38 (or 28), the difference 936
937 is 5 and the answer is considered correct. But if you respond with 41 or more (or 25 or 937
938 less), for example, the difference would be greater than 7 and the answer is incorrect. 938

939 Please move the sphere to indicate which percentile you think your score falls in: 939

940 [Slider with values from 0 to 100] 940

941

 941

942 **Recommendation**

942

943 Among the people with whom you have taken any class at the university, who is your 943
944 recommendation for the [Subject Area] test? Please write that person's name in the 944
945 box below: 945

946 **Important:** You will not be considered for payment unless the recommended 946
947 person is someone with whom you have taken at least one class during your 947
948 studies. 948

949 Your response is only a recommendation for the purposes of this study and we will **not** 949
950 contact your recommended person at any time. 950

951 [Clicking shows the explanations below] 951

952 Who can I recommend? 952

953 Your recommendation **must** be someone with whom you have taken (or are taking) a 953
954 class. If not, your answer will not be considered for payment. The person you recommend 954
955 will not be contacted or receive any benefit from your recommendation. 955

956 As you write, you will see up to 7 suggested student names containing the letters you 956
957 have entered. The more you write, the more accurate the suggestions will be. Please 957
958 write **without** accents and use the letter 'n' instead of 'ñ'. If the name of the person 958
959 you're writing doesn't appear, it could be because you made an error while writing the 959
960 name. 960

961 If the name is correct and still doesn't appear, it could be because the student is not en- 961
962 rolled this semester or didn't take the SABER 11 test. In that case, you must recommend 962
963 someone else. 963

964 My earnings for this question? 964

965 For your recommendation, you could receive earnings of up to 250,000 (two hundred and 965
966 fifty thousand) PESOS. The earnings are calculated based on your recommendation's 966
967 score and the percentile of that score compared to other UNAB students, as follows: 967

- 968 • We will multiply your recommendation's score by \$100 (one hundred) pesos if it's 968
969 between the 1st and 50th percentiles 969
- 970 • We will multiply your recommendation's score by \$500 (five hundred) pesos if it's 970
971 between the 51st and 65th percentiles 971
- 972 • We will multiply your recommendation's score by \$1000 (one thousand) pesos if 972
973 it's between the 66th and 80th percentiles 973
- 974 • We will multiply your recommendation's score by \$1500 (one thousand five hun- 974
975 dred) pesos if it's between the 81st and 90th percentiles 975
- 976 • We will multiply your recommendation's score by \$2500 (two thousand five hun- 976
977 dred) pesos if it's between the 91st and 100th percentiles 977

978 This is illustrated in the image below: 978

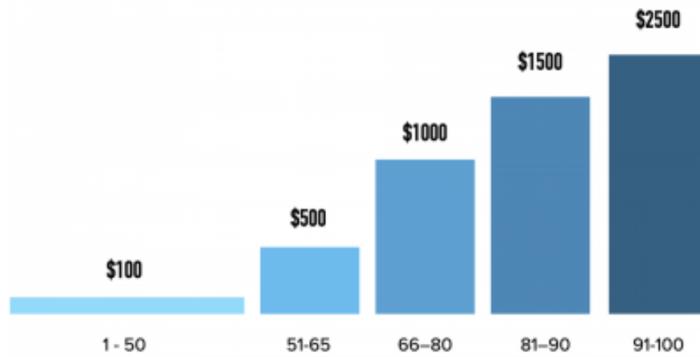


Figure B.1: Earnings for recommendation questions

979 For example, if your recommendation got 54 points and the score is in the 48th percentile, 979

980 you could earn $54 \times 100 = 5400$ PESOS. But, if the same score of 54 points were in the 980
981 98th percentile, you could earn $54 \times 2500 = 135,000$ PESOS. 981

982 [Text field with student name suggestions popping up as participant types] 982

983 _____ 983

984 **Relationship with your recommendation** 984

985 How close is your relationship with your recommendedation: “[Name of the student 985
986 selected from earlier]”? (0 indicates you are barely acquaintances and 10 means you are 986
987 very close) 987

988 [Slider with values from 0 to 10] 988

989 _____ 989

990 **Your recommendation’s score** 990

991 Compared to university students, in which percentile do you think [Name of the student 991
992 selected from earlier]’s score falls in the **[Subject Area]** test (1 is the lowest percentile 992
993 and 100 the highest)? 993

994 [Clicking shows the explanations below] 994

995 How is a percentile calculated? 995

996 A percentile is a position measurement. To calculate it, we take the test scores for all 996
997 students currently enrolled in the university and order them from lowest to highest. The 997
998 percentile value you choose refers to the percentage of students whose score is below 998
999 yours. For example, if you choose the 20th percentile, you’re indicating that 20% of 999
1000 students have a score lower than yours and the remaining 80% have a score higher than 1000
1001 yours. 1001

1002 What can I earn for this question?

1002

1003 For your answer, you can earn **20,000 (twenty thousand) PESOS**, but only if the 1003
1004 difference between your response and the correct percentile is less than 7. For example, 1004
1005 if the percentile where your recommended person's score falls is 33 and you respond with 1005
1006 38 (or 28), the difference is 5 and the answer is considered correct. But if you respond 1006
1007 with 41 or more (or 25 or less), for example, the difference would be greater than 7 and 1007
1008 the answer is incorrect. 1008

1009 Please move the sphere to indicate which percentile you think your recommended per- 1009
1010 son's score falls in: 1010

1011 [Slider with values from 0 to 100] 1011

1012 _____ 1012

1013 Demographic Information 1013

1014 What is the highest level of education achieved by your father? 1014

1015 [Primary, High School, University, Graduate Studies, Not Applicable] 1015

1016 What is the highest level of education achieved by your mother? 1016

1017 [Primary, High School, University, Graduate Studies, Not Applicable] 1017

1018 Please indicate the socio-economic group to which your family belongs: 1018

1019 [Group A (Strata 1 or 2), Group B (Strata 3 or 4), Group C (Strata 5 or 6)] 1019

1020 _____ 1020

1021 UNAB Students Distribution

1021

1022 Thinking about UNAB students, in your opinion, what percentage belongs to each socio- 1022
1023 economic group? The total must sum to 100%: 1023

1024 [Group A (Strata 1 or 2) percentage input area] 1024
1025 [Group B (Strata 3 or 4) percentage input area] 1025
1026 [Group C (Strata 5 or 6) percentage input area] 1026
1027 [Shows sum of above percentages] 1027

1028 ————— 1028

1029 End of the Experiment

1029

1030 Thank you for participating in this study. 1030

1031 If you are chosen to receive payment for your participation, you will receive a confirma- 1031
1032 tion to your UNAB email and a link to fill out a form with your information. The process 1032
1033 of processing payments is done through Nequi and takes approximately 15 business days, 1033
1034 counted from the day of your participation. 1034

1035 [Clicking shows the explanations below] 1035

1036 Who gets paid and how is it decided? 1036

1037 The computer will randomly select one out of every ten participants in this study to be 1037
1038 paid for their decisions. 1038

1039 For selected individuals, the computer will randomly select one area: mathematics, 1039
1040 reading, or English, and from that area will select one of the belief questions. If the 1040
1041 answer to that question is correct, the participant will receive 20,000 pesos. 1041

1042 The computer will randomly select an area (mathematics, critical reading, or English) to 1042
1043 pay for one of the recommendation questions. The area chosen for the recommendation 1043
1044 question is independent of the area chosen for the belief question. The computer will 1044
1045 take one of the two recommendations you have made for the chosen area. Depending on 1045
1046 your recommendation's score, you could win up to 250,000 pesos. 1046

1047 Additionally, people selected to receive payment for their participation will have a fixed 1047
1048 earnings of 70,000 pesos for completing the study. 1048

1049 _____ 1049

1050 **Participation** 1050

1051 In the future, we will conduct studies similar to this one where people can earn money 1051
1052 for their participation. The participation in these studies is by invitation only. Please 1052
1053 indicate if you are interested in being invited to other studies similar to this one: 1053

1054 [Yes, No] 1054