The Effect of College Desegregation on Academic Achievement and Students' Social Interactions: Evidence from Turnstile Data

Tatiana Velasco Rodriguez[*]

October 28, 2021

[Most recent version here](#)

Abstract

Efforts to desegregate schools by helping low–income students attend elite institutions have spread around the world. However, the benefits of socioeconomic integration may fail to emerge if social interactions within schools remain segregated. In this paper, I study a natural experiment at an elite university that experienced a sharp and unexpected increase in its enrollment of low—income students, and use it to measure the effect of desegregation policies on students' social interactions and academic achievement. To identify students' interactions, I develop a measure based on students' co-movements across campus as recorded by turnstiles guarding all entrances. Increasing exposure to desegregation led to the diversification of students' social interactions with no adverse effects on academic achievement. Moreover, I find at least half of the increase in interactions between wealthy and low–income students is explained by interactions with low–income high–achieving students, which is consistent with a model where wealthy students links other low–income students by avoiding adverse peer effects.

1 Introduction

The separation of students by socio–economic status, race, or ethnicity is pervasive in education. To address the issue, policymakers have implemented desegregation policies such as financial aid and affirmative action programs that reallocate low–income and underrepresented minorities into selective institutions. Prior research has extensively documented the positive impacts that attending elite institutions has on low–income and underrepresented students (Chetty et al., 2020; Dale and Krueger, 2002; Hoekstra, 2009; Londoño-Velez et al., 2020). While there is evidence suggesting social connections matter in explaining the impacts of elite institutions (Zimmerman, 2019; Michelman et al., 2020; Rao, 2019), research also suggest changes in peers composition can lead to segregation within schools (Carrell et al., 2013). Moreover, desegregation can lower the quality of peer effects, ultimately harming the achievement of traditionally privileged students attending these institutions (Arcidiacono et al., 2015). In this paper I ask, can college desegregation diversify social interactions without harming academic performance?

To study this, I make use of a natural experiment at a large elite college in Colombia which experienced a sharp and unexpected increase in the enrollment of low–income students, driven by the introduction of a nationwide desegregation program known as *Ser Pilo Paga* (SPP). To measure social interactions, I assemble a novel database of over a hundred million records of students' movements across campus as recorded by turnstiles guarding all campus entrances. I develop a measure to identify which students socialize with one another based on how commonly I observed them entering and exiting campus buildings together, and I validate it against a survey where students listed their friends and acquaintances. I combine these data with student–level records on course enrollment and compute academic achievement and persistence measures. I find that the increased exposure to low–income students significantly increases the interactions between wealthy and low–income peers, with no adverse effects on the achievement of the wealthy students.

My empirical analysis leverages the plausibly random variation in the exposure to low–

income peers within degree majors and across entry cohorts enrolling before and after the introduction of SPP. Specifically, I use a difference–in–differences approach that exploits the variation in the percentage of low–income peers within majors and across entry cohorts, and I measure the effects of exposure to desegregation on the academic achievement, persistence, and diversity of social interactions of wealthy students traditionally attending this university. I conduct several test showing the variation in exposure is plausibly random, and that the measurement error in social interactions does not lead to bias in my estimates

This paper starts by describing the policy context, data, and research design used for the analysis. In October of 2014, the Colombian government launched *Ser Pilo Paga*, a policy that targeted low-income students with outstanding academic achievement to promote their attendance to high-quality universities in the country. The program consisted of a loan that covered 100 percent of the tuition cost plus a small stipend. The loan was forgiven upon completion of the degree. *Ser Pilo Paga* induced an influx in the number of low-income students enrolled at several high-quality private universities in the country, closing the socio–economic enrollment gap among high achievers (Londoño-Velez et al., 2020). This paper focuses on one large private university, where the number of low-income students enrolled unexpectedly tripled in the Spring of 2015. Because of the short timing of the policy, universities and wealthy students had no time to adjust their application and admission criteria in ways driven by their preferences for low-income peers. Thus, the policy induced plausibly random variation in the share of low–income peers across majors and entry cohorts, which I use to capture wealthy students' exposure to low–income peers. I conduct numerous tests to show there is no students' selection in majors or entry cohorts and no unobserved exposure through courses that confounds my findings. More importantly, I show that measurement error on the turnstile–elicited interactions outcomes do not bias the estimated effects.

Next, I discuss the estimations of the effects of exposure to desegregation on academic achievement and persistence in college. I find a positive yet modest impact of exposure on the number of credits attempted by the first and third term of college, which did not

persist to the sixth term, and I do not find impacts on GPA. The average increase in the percentage of low–income peers was 9.51 percentage points, which led to 0.02 and 0.1 more credits attempted by wealthy students by the first and third term, respectively. Both of these are equivalent to 0.04 standard deviations of the pre–SPP distribution. Regarding college persistence, I find a positive but small impact on the probability of dropout of 0.2 percentage points – equivalent to 0.06 standard deviations from the pre–SPP dropout rate distribution. This impact does not map with changes in the probability of graduation in eight terms (graduation on time). Plus, I do not find evidence of differential impacts for students exposed to shares of low–income peers at the top 25 percent of the distribution (i.e., > 30 percent of low–income peers), or evidence suggesting adverse peer effects due to the on average low–achievement of low–income students.

My findings on the effects of exposure to desegregation indicate modest to null impacts on achievement and persistence. Such results could be explained by a lack of interactions between wealthy and low–income students. Wealthy students could avoid interactions with low–income peers if they anticipate adverse peer effects on achievement. Simultaneously, low–income students may prefer to keep their interactions among themselves, a typical behavior known as *homophily*.

The second part of my empirical analysis examines whether changes in exposure to the low–income changed the diversity of social interactions. I examine the effect that the increased exposure to low–income peers had on the number of links between wealthy and low–income students and on the probability of a wealthy student forming a link with a low–income peer. The average increase in the percentage of low–income peers of 9.51 percentage points augmented the links between wealthy and low–income peers by 0.28 and increased the probability of a wealthy student having any link with a low–income by 0.7 percentage points. Relative to the distribution of these variables among the pre–SPP entry cohort, this represents an increase of 0.51 standard deviations in the number of links and an increase in 0.2 standard deviations in the probability of a link with a low–income student. Notably,

the increased exposure to low–income peers led to reductions in the number and probability of links among wealthy students, albeit the size of the effects was small (0.08 and 0.005 standard deviations, respectively).

Next, I examine the effect of exposure to low–income peers on the overall composition of wealthy students' social connections. The 9.5 percentage points increase in the percentage of low–income peers translates in a 6.8 percentage points increase in the percentage of wealthy students' links with the low income (equivalent to 0.6 standard deviations on the pre–SPP cohort). Importantly, I find large and significantly positive effects on the number and the probability of links with low–income peers for students exposed to shares of low–income peers at the top 25 percent of the distribution, suggesting the positive impacts of exposure on the diversity of interactions between wealthy and low–income students persist even among the groups with the most considerable changes in the composition of students.

Lastly, I test the theoretical prediction of wealthy students diversifying their interactions primarily with low–income high achieving students from whom they do not anticipate adverse peer effects on achievement. My results suggest about half of the impact of exposure on the number of links with the low–income peers is explained by links with the high achievers. Coupled with the modest impacts on the number of credits attempted, my results suggest potentially positive effects from interactions with the low–income high–achieving students, particularly when achievement is measured by the number of credits attempted.

This study makes three contributions to the literature. First, my results on the effects of desegregation on achievement and persistence contribute to the literature on the effects of desegregation on academic outcomes by providing evidence from a selective higher education setting. Except for Arcidiacono and Vigdor (2010) who find adverse effects, the evidence on the impacts of desegregation on the academic achievement of college students is relatively scarce.[1] My findings expand these literature and dispute Arcidiacono and Vigdor (2010) by

---

[1]Arcidiacono and Vigdor (2010) use quasi-random variation across entry cohorts at selective universities in the US to study the effects of exposure to minorities on white and Asian students' achievement, finding negative effects.

showing evidence that increases in the exposure to underrepresented students at elite schools driven by affirmative action policies have zero effect on the privileged students traditionally attending selective colleges and, if anything, can lead to modest improvements in early outcomes. In that regard, my findings align with previous evidence from K–12 settings which have found no effect of desegregation policies on the academic achievement of students traditionally attending these institutions (Angrist and Lang, 2004; Dobbie and Fryer, 2014).[2] A related branch of this literature has examined the effect of desegregation and affirmative action on students behaviors like pro–social behaviors and political affiliations (Rao, 2019; Boisjoly et al., 2006; Billings et al., 2021; Londoño-Vélez, 2020).[3] A common aspect of this literature is that social interactions are pointed as a potential channel explaining the result. My work provides direct evidence showing that social interactions diversify under desegregation exposure.

Second, my study is the first to estimate how students' social interactions at elite colleges change at the outset of financial aid and affirmative action policies fostering desegregation. While prior research has consistently found positive impacts on the college attainment of underrepresented students benefiting from financial aid and affirmative action programs (Bleemer, 2020; Chetty et al., 2020; Londoño-Velez et al., 2020; Mello, 2021), the examination of the role of social interactions in explaining these findings has been somewhat elusive. Simultaneously, research examining how elite colleges foster social mobility suggests membership to elitist social groups such as elite high schools or clubs in college is a channel explaining the lack of long–term social mobility and success for low–income students at these institutions (Michelman et al., 2020; Zimmerman, 2019). In this paper, I develop a

---

[2]Angrist and Lang (2004) studied the effect of a desegregation program in Boston on the academic achievement of the students traditionally attending the receiving schools, finding no significant impact; a similar study by Dobbie and Fryer (2014) focuses on students eligible to attend schools with high achieving peers and finds no impacts on the achievement of either group.

[3]My work is closely aligned to that of Londoño-Vélez (2020), who studied the effect of socio-economic diversity at an elite college in Colombia on students' redistribute preferences. In this work, Londoño-Vélez finds positive impacts of exposure on wealthy students' preferences - a result that seems to be related to more interactions with low-income peers. My work validates the latter finding while pointing out that the change in social interactions is relatively small.

robust measure of social interactions to show the likelihood and extent to which social interactions among students from different socioeconomic backgrounds increase at elite colleges participating in desegregation policies.

Third, my findings also contribute to the literature examining diversity in schooling settings and its effects on segregation in social networks. This research has examined the process under which friendships form in college settings and has relied on proxies of social interactions such as email exchanges (Marmaros and Sacerdote, 2006) or Facebook friendships (Baker et al., 2011). Evidence coincides in that peers' proximity and peers' race are determinants of friendship formation. Namely, students assigned to the same dorm are more likely to be connected, but the chances are higher for same–race students.[4] My study uses a different dimension of proximity which is being in the same major and entry–cohort. My findings indicate that proximity through majors and cohorts group is determinant for students interactions. A related sub–stream of research has focused on measuring overall segregation in social interaction and on studying how policies can reduce within-group segregation in K–12 settings, finding no association between who students interact with and academic achievement (Echenique et al., 2006), and finding non–linear responses in interactions to scenarios of minorities reallocation across schools (Mele, 2020).[5] My findings show consistently positive impacts on the diversity of interactions along different levels of exposure and show that changes in interactions through changes in exposure do not lead to impacts on academic achievement.

---

[4]Marmaros and Sacerdote (2006) examine how people form social networks with their peers. They use emails exchange data from students and find that first-year students form friendships with students in the proximity and are more likely to form friendships with peers of the same race. Baker et al. (2011) use data from Facebook and random dorm assignment at one college and finds exposure to different races via dorms leads to more diverse friendships.

[5](Echenique et al., 2006) measure within–school segregation as the extent to which students interact socially with other students from the same race. Mele (2017) develops a structural model of friendship formation among students, and Mele (2020) use it to simulate reallocation programs across schools and examine its impacts on within school friendship formation. His findings suggest that policies that reallocate students by parental income have less impact on racial segregation within schools.

2 Conceptual Framework

Desegregation policies in higher education have shown to be effective at improving the academic achievement of minorities and under–represented students, while fostering positive views towards others among wealthy students (Bleemer, 2020; Mello, 2021; Londoño-Velez et al., 2020; Castleman and Long, 2016; Fack and Grenet, 2015; Dale and Krueger, 2002; Hoekstra, 2009; Long, 2008; Chetty et al., 2020; Canaan and Mouganie, 2018; Rao, 2019; Boisjoly et al., 2006; Londoño-Vélez, 2020). While social interactions are a plausible channel driving positive effects, prior evidence suggests segregation within schools may be persistent and even fostered under desegregation policies.

Students traditionally attending selective colleges may experience negative peer effects if the overall level of their peers' performance diminishes (Arcidiacono et al., 2015; Arcidiacono and Vigdor, 2010). Furthermore, wealthy students could avoid interactions with the incoming low achieving peers if they anticipate adverse effects on their own achievement (Carrell et al., 2013). Similarly, elite college students may segregate through group memberships such as high school background, social clubs, and Greek–letter organizations. These forms of segregation can have positive payoffs in the job search by providing exclusive information on job referrals and social and cultural capital that signals value to job recruiters and facilitates placement in better-paying jobs (Marmaros and Sacerdote, 2002; Zimmerman, 2019; Michelman et al., 2020) (Rivera, 2016). Equally important, wealthy and low-income students can tend to segregate simply by virtue of their socialization choices. Students may prefer to interact with others with whom they share certain characteristics (a.k.a. homophily) (Christakis et al., 2010; Currarini et al., 2009), and the cost of meeting new people can be lower if students have already friends in common (Rogers and Jackson, 2007).

The prediction derived from this literature is that exposure to desegregation within an elite college can have no impact on the interactions between wealthy and low–income students if wealthy students anticipate those can diminish their academic performance and the value of their information and social capital in the future job search. If wealthy students interact

with the low–income, then those must be with low-income students who are on similar levels of achievement to them and therefore cannot harm their performance. At the same time, low-income students can be less prompted to form interactions with the wealthy if they have a larger pool of other low-income peers to form interactions with, a behavior consistent with homophily. As a result, if social interactions do diversity, then it must be only with low–income students who are high achievers, such that the academic achievement of wealthy students is not harmed and, if anything, is marginally improved.

3 Policy Context

In this project, I examine the effect of a socio–economic desegregation policy on students' academic achievement and social interactions. Specifically, I study the case of a large private university located in Bogotá, Colombia (from now on *Elite University*[6]), which in 2015 experienced a large and unexpected increase in the number of low–income students enrolled, while keeping the enrollment of relatively wealthy students constant. The increase was driven by *Ser Pilo Paga* (SPP) – a forgivable loan program for high–achieving low–income students who wished to attend a high-quality university. Importantly, the increased enrollment of low–income students varied across the thirty-one degree majors offered at Elite University. In my research design, I focus on relatively wealthy students and compare students from the entry cohorts before and after SPP (2014 vs. 2015). I use the change in the number of low–SES students across degree majors as the treatment. In this section, I explain the context of SPP and Elite University where the natural experiment took place.

Higher education in Colombia is strongly segregated by socio–economic background. High–quality private universities have exceedingly expensive tuition rates relative to average salaries in the country; supply of public university seats is stagnant; and financial aid is scant (Marta Ferreyra et al. 2017). These factors led students to sort across colleges by

---

[6]This is a made-up name. I do not provide the real name of the university I study for confidentiality reasons.

socio-economic status (Camacho et al., 2017). SPP aimed to combat this segregation by providing low-income students a loan that covered tuition plus a small allowance for attending a high-quality accredited institution.[7] The loan was forgiven conditional on completion of the degree. Eligibility to SPP required that students were classified as poor under the governments' index of household wealth, and scored in the top ten percentile of the national high school exit exam SABER 11.[8] SPP awarded loans for new cohorts of students between 2015 and 2018 benefiting about 40,000 students nationwide. Previous research has found SPP increased diversity at top private universities by making the selection mechanism based more on ability than on income (Londoño-Velez et al., 2020).

The timing of SPP and the admission rules at Elite University set the conditions of the natural experiment I exploit in my research design. First, admissions to Elite University are open for the Spring and Fall term of each year and are determined by the applicant's score in the SABER 11 standardized test. Students must apply to a degree major[9] and entry cohort for which admission officers had pre–determined a specific SABER 11 weighting formula[10] and cutoff score. Second, SPP was widely unexpected by students and higher education institutions. SPP was launched in October of 2014 and only students who had taken that October's test were eligible. Candidates had to apply for enrollment in the following Spring of 2015, for which 10,000 forgivable loans were offered. Thus, students who traditionally applied to Elite University had very little time to change their application portfolio and university officers could not adjust the admission criteria to limit the influx of admitted and

---

[7]The high-quality accreditation is granted to higher education institutions by the National Council of Accreditation. It is granted after a detailed review from a panel formed by the Institution, the academic community, and the Council. By 2014, the year of the first round of SPP, 32 universities in Colombia had high–quality accreditation.

[8]The household's index of wealth is known as SISBEN and it is based on the census survey targeted to household previously screened as potentially poor. Londoño-Velez et al. (2020) provide more details about how SISBEN was used to screen SPP eligible students. SABER 11 is a requirement for all students in the country who are about to complete their high school education. The exam is applied twice a year, following the two academic calendar of schools in the country: January – November and August – June.

[9]As opposed to the U.S., applicants to higher education must apply to a major for degree as well to as a college.

[10]The SABER 11 is made of five modules which are given different weights depending on the major of application. For example, for admission to engineering majors, quantitative reasoning is a assigned a higher weight than the social sciences module

eventually enrolled students. As a result, the number of middle– and high–income students enrolled in 2015 remained similar to that from 2014, but the number of low-income students increased significantly.

Figure 1 depicts the first–term enrollment trends by socio–economic status (SES) at Elite University. Between 2012 and 2014, less than 150 first–term students came from low–SES backgrounds. Once the first cohort of SPP beneficiaries enrolled, the number of low–SES students tripled to 541, while the number of students from other socio–economic backgrounds remained almost the same. Figure 2 compares the number of low–SES students across majors, in the entry cohorts before and after SPP. Red bars depict the number of low–income students in the cohort right before SPP (i.e., 2014), whereas blue bars depict the number of low–income students in the first cohort of SPP (i.e., 2015). The variation in the number of low–SES students is important. Majors such as Business and Music experienced virtually no change in the number of low–SES students, while others like Civil Engineering or Psychology experienced a notable increase.

To examine whether SPP led to crowding out of wealthy students by new incoming low–income peers, I plot in the secondary axis of Figure 2 the percentage change in the number of wealthy students (i.e., Middle– and High–SES) in 2015, relative to the 2014 cohort, and for each major. If the increase in the number of low–income students across majors had led to crowing out of the wealthy students traditionally attending these programs, then the percentage change in the number of wealthy students should decrease as the number of low–income students increases. Similarly, majors with virtually no change in the number of low–income students should show no percentage change in the number of wealthy students. However, neither of those is the case. Visual inspection of Figure 2 shows that the percentage change in the number of wealthy students enrolled in each major from 2014 to 2015 is not associated with the increase in the number of low–income students.

4   Data and Descriptive Statistics

The data for this project comes from two sources: administrative records from Elite University, and detailed records from turnstiles located in each of the 18 access points to Elite University campus.

Elite University administrative records. I use records from all students enrolled at Elite University between 2012 and 2018 which contained student-course level data on student characteristics (i.e. sex, age, mother's education, High School ID), SABER 11 standardized test scores, SPP recipient status, selected degree major, entry cohort and term of enrollment. For each semester, I observe each of the courses in which the student is enrolled and their course GPA. More importantly, I observe the student's household social strata indicator. This indicator has six categories which are used to provide homes with subsidies in utility bills. Plus, it is also widely known in the country as a proxy of social status. I use the household social strata to classify students in three socio-economic status (SES): middle- and high-SES – which I will refer to as relatively wealthy students, and low–income students. low–income are students from strata one and two, middle-SES are students from strata three and four, and high-SES are students from strata five and six. Students benefiting from SPP mostly fall in the low–income category. As depicted in Figure 1, the majority of students at Elite University are classified as high- and middle-SES.

Turnstile records. I use records on student access and exits to Elite University campus to identify students' social interactions. Elite University campus is guarded by turnstiles located at the 18 entrances to main buildings and campus areas. In order to enter or exit through any of these entrances students and university staff must tap their University ID. Security officers at Elite University provided me individual-level records of University ID taps on the turnstiles from February 1st, 2016 to November 1st, 2019. These records include student ID number, entrance, action (IN or OUT of campus), and the date, hour, minute and second of the tap. Figure 3 displays a heat map of the average frequency of student ID taps at each of the 18 entrances to campus by 20 minutes blocks and for each academic term since

the Spring of 2016. Yellow cells and blue cells indicate peak and off-peak hours respectively. The figure documents the constant flow of students across the campus entrances throughout the day, with peak hours at times of class change as well as during lunch hours.

I define a pair of students as linked when their IDs are tapped at a turnstile in a time window of three seconds or less, in the same entrance and direction (either entering or exiting campus), and when I observed the same pair of IDs co-moving at least twice in a semester. Appendix A describes the data validation process for this definition.

Sample. My analytic sample consist of all the first–term students in the entry cohorts before and after SPP (i.e. Fall and Spring of 2014 and 2015). I search for their interactions during the 6th and 7th calendar semesters after their first–term of enrollment, and among students in the same entry–cohort and degree major. For example, I match students in the entry cohort of Spring of 2014 with their interactions as captured by the turnstiles during the Fall of 2016 and the Spring of 2017. I merge administrative records and pairwise–level students' interactions data using the student ID number which is available in both data sources. My final sample consist of 5,955 students across 31 degree majors and 4 entry cohorts. This sample captures the universe of students enrolled in these majors and cohorts except for two majors (Government, and the Directed Studies major) which started after SPP.

Student characteristics:. Table 1 provides descriptive statistics of wealthy and low–income students in the pre– and post–SPP entry cohorts (i.e., 2014 vs. 2015). I divide the sample between relatively wealthy students - that is, middle-SES and high—SES students, and low–income students, before and after the implementation of SPP (i.e., 2014 vs. 2015). The table includes the mean differences between low–income and wealthy students. In 2015, 81 percent of low–income students at Elite University were SPP recipients. About half of wealthy students are middle–SES students. In both 2014 and 2015 cohorts, wealthy students are more likely to be females, are slightly older and with mothers more educated than low–income students. Also, wealthy students have higher SB11 test scores than low–income ones,

12

and the gap increases and becomes statistically significant among the 2015 cohort. The gap in SB11 test scores between wealthy and low–income students was 0.10 standard deviations in 2014, but increased to 0.26 standard deviations in 2015.

Wealthy students have on average more links than low–income peers with others in their major and entry cohort (5.21 vs. 4.94 links in 2014), and the difference becomes statistically significant among the 2015 cohort (5.53 vs. 4.60 links). Before SPP, the number of links with other low–income students is statistically the same among wealthy and low–income students (0.24 vs. 0.35, respectively). But, in 2015, wealthy students have significantly fewer links with other low–income than their low–income peers (1.73 vs. 0.59, respectively). Importantly, wealthy students have on average more peers from high school enrolling at Elite University in their same cohort than low–income students in both pre– and post–SPP cohorts (11.54 vs. 3.17 in the 2014 and 11.73 and 1.98 in the 2015). There are no statistically significant differences between the number of ID taps at the turnstiles of wealthy and low–income students for either the 2014 or the 2015 entry cohorts.

Table 1 also describes the average characteristics of the links of both wealthy and low–income students. The characteristics of the links of wealthy and low–income students were statistically the same among the students in the 2014 cohort, except for the share of links from the same high school, which is larger among wealthy students (0.04 vs. 0.01). But in the 2015 cohort, wealthy students exhibit a larger share of same–gender links than low–income students (0.51 vs. 0.45 links), and have a larger SB11 test score difference with their links (0.79 vs. 0.67). The latter suggest wealthy students enrolling in 2015 exhibit less diversity in their social interactions in terms of certain demographics, but more in terms of pre–college academic achievement, relative to the low–income peers.

Students' academic achievement: I characterize the differences in academic achievement between wealthy and low–income students in Figure 4 and Figure 5. For these figures, I take advantage to the administrative data availability and plot the trends in academic achievement across the entry cohorts starting with 2012. For each cohort, I plot the average

achievement outcome among wealthy and low–income students, and include the estimated 95 percent confidence interval based on clustered standard errors at the major and entry cohort level. The red line separates the entry cohorts before the start of SPP (left side) and the cohorts entering during SPP (right side). Figure 4 displays performance indicators, mainly cumulative GPA and total credits attempted, whereas Figure 5 describes persistence (dropout rates and graduation). I label a student as a dropout if they do not show up as enrolled during two consecutive terms after their fifth term of college. Similarly, I label a student as graduated if they completed their degree in eight terms or less. At Elite university, this is considered as graduation on time for all their degrees except medicine.

The cohort of wealthy and low–income students that enrolled Elite University at the outset of SPP (i.e., the 2015 entry cohort) exhibit significant achievement gaps, particularly in their GPA and cumulative credits attempted, with low–income students having on average lower cumulative GPA and fewer attempted credits than their wealthy peers. For example, the GPA of pre–SPP cohorts is relatively constant and close to 3.85 for both wealthy and low–income students. But for the SPP cohort, the GPA of low–income students drops to 3.75 in their first term of college and to 3.6 by their third term of college, while the GPA of wealthy students remains the same. Regarding the cumulative number of credits attempted the pre–SPP cohorts of wealthy and low–income students have attempted on average 50 and 48 credits by the third term, respectively. But in the SPP cohort, low–income students have on average attempted 45.7 credits while wealthy students attempted the same number of credits. A course at Elite University usually bears 3 credits. This means that low–income students enrolling in 2015 had attempted on average at least one class less than their wealthy students peers by the third term of college, and with a cumulative GPA that is 0.25 below that of their wealthy peers. Nevertheless, the differences in achievement did not paired with differences in dropout or graduation rates, suggesting the relatively low achievement of low–income students did not translate in diminished persistence.[11]

---

[11]Importantly, graduation rates in less than eight terms are very small at Elite University across all groups as many students tend to take extra semesters to course minor degrees or to double major with other degree.

5 Identification Strategy

In this paper, I use a difference–in–differences strategy to examine the impact of increased exposure to low/-income students' academic achievement and social interactions. In particular, I focus on the impacts on wealthy students, and I exploit the variations in the percentage of low–income students enrolling in each cohort and across the different majors at Elite University. Specifically, I estimate:

$$Outcome_i^{mc} = \beta_l R_{mc}^l + \mathbf{X_i'}B + \beta_m + \beta_c + \varepsilon_{imc} \tag{1}$$

In equation 1 $Outcome_i^{mc}$ represents either the academic achievement outcome of a relatively wealthy student i enrolled in majors m and entry cohort c or the number of links of the student with other low–income peers. R_{mc}^l is the percentage of low–income students in student i majors and entry cohort i.e., $R_{mc}^l = \frac{N_{mc}^l}{N_{mc}} * 100$ where N_{mc}^l is the number of low–income students and N_{mc} is the total number of students enrolled in student i major m and entry cohort c. Figure 2 describes the variation exploited for causal identification. Namely, the percentage of low–income students R_{mc}^l across majors m increased for the 2015 entry cohorts and relative to the 2014 entry cohorts at different rates across majors, without leading to crowding out of wealthy students. Equation 1 includes controls for student i characteristics that can determine the relation between the exposure to low–income students, and achievement and social interactions. Specifically, $\mathbf{X_i}$ is a matrix of female indicators, age in years at the start of college, Mother's with no college education, SB11 score, and middle–SES indicators. β_m and β_c capture major and entry cohort fixed effects, which absorb unobserved variation common to majors and to entry cohorts, respectively. Finally, ε_{imc} represents robust standard errors clustered at the major and entry cohort levels. I estimate Equation 1 using Ordinary Least Squares.

---

low–income students benefiting from SPP and other financial aid programs tend to be constrained in that they do not get financed for terms beyond those scheduled in their major curriculum, which explains their slightly higher likelihood of graduation.

Identification of β_l relies on three assumptions: in the absence of the treatment, the outcomes exhibit the same trends for the treated and control groups (i.e., the parallel trends assumptions), there is no self-selection in exposure to low–income students within major-cohorts groups of students (i.e., unobserved exposure effects), and the measurement error in the turnstile–elicited links do not contain non–random variation confounding the effect of exposure to the low–income (i.e., measurement error in turnstile–elicited interactions). Table 2 summarizes the identification challenges and the exercises I conducted to assess them. The table conducts to the Figures and Appendix sections in which I provide details of the analyses conducted.

The parallel–trends assumption implies that any differences in the outcomes of wealthy students across majors and within entry cohorts can be attributed only to changes in exposure to low–income peers, thus ensuring causality. Violations to this assumption imply there is non–random allocation of wealthy students across majors and entry cohorts driven by the change in the percentage of low–income peers. Specifically, at the outset of SPP, wealthy students could have self–selected in majors and entry cohorts due to their preferences for low–income peers. Alternatively, the influx in low–income students could have crowed out wealthy ones. I examine the trends in the outcomes and student characteristics using an event study to test for this.[12] Results are displayed in Figure 6 for student characteristics, and in Figure 7 and 8 for student academic achievement and persistence indicators. The estimated effects are the same across all the pre– and post–SPP entry cohorts. Thus, the results suggest that the allocation of wealthy students across majors and cohorts did not change, and its relation with the percentage of low–income peers in the group did not change with SPP. Coupled with the tight timing of the policy, which gave Elite University no time to adjust admission criteria in ways that would crow out students based on their eligibility

---

[12]Namely, I estimate the following equation for both outcomes and student characteristics:

$$Y_{imc} = \sum_{c=2012}^{c=C} \mu_{lC} R^l_{m,c=C} + \mathbf{X'_i}M + \mu_m + \mu_c + \varepsilon_{imc}$$

16

for admission (see Figure 2), I conclude there is no evidence indicating the parallel–trends assumption is violated in this context.

The second source of bias deals with unobserved exposure to low–income peers due to the composition of first-term courses taken by each wealthy student in a major and cohort. Exposure to low–income students operates not only through majors and cohorts but also in the classroom. Several of the courses offered to first–term students are open to multiple majors, which may lead to variations in exposure to low–income peers not accounted for by my initial difference–in–differences research design. For example, students at all the Engineering, Economics, and Business programs must take Differential Calculus in their first term. As a result, students in Business majors –which had virtually no change in the number of low–income students enrolled, may take Calculus with low–income peers from other majors, thus being exposed to low–income students in ways not accounted for by the entry cohort and major variation captured by R^l_{mc}. If students within the same major and entry cohort deferentially self–select to take specific first–term courses based on preferences for low–income peers, the estimated β_l could be biased even after accounting for unobserved variation common to majors and cohorts. I test for the presence of this bias by exploiting the within major–cohort variation in the number of low–income students at each of the wealthy students' first term courses. To ensure exogeneity, I instrument the number of low–income peers by its predicted allocation across courses based on courses enrollment data from 2012 and 2013. Appendix B describes the methodological details and results from this test. Overall, I do not find evidence that student selection of first–term courses is biasing my Difference–in–Differences estimates.

The last bias concern is measurement error in the turnstile–elicited interactions and the risk of capturing random co–movements across the turnstiles, thus falsely depicting the effects of exposure to low–income on students' social interactions. If turnstile–elicited interactions partially capture true social interactions, then those need to be on average representative of true interactions and cannot be biased due to the potential random noise

17

in the measurement error. Moreover, the rate of false–positives and false–negatives (i.e., the likelihood of defining a pair of students as linked when in fact they are not; and the likelihood of defining a pair as not linked when in fact they are, respectively) cannot be determined by the exposure to low–income students R_{mc}^l. My definition of students' social links accounts for these possibilities and aims to minimize the rate of false–positives and negatives. Specifically, I use secondary data on the survey–elicited social interactions among one major–cohort group at Elite University to obtain estimates of the rates of false–positives and false–negatives under alternative turnstile–elicited links definitions.[13]

Appendix A provides details on the secondary data and the computation framework and procedures I use to assess measurement error. I find that turnstile–elicited links suffer a relatively large rate of false–negatives of approximately 60 percent, but a rate of false– positives below the 10 percent (see Table 7). To assess the extent to which measurement error can diminish the quality of the turnstile–elicited interactions, I compare the average characteristics of turnstile–elicited links with those from survey–elicited links and with those obtained under a simulated scenario of turnstile–elicited links formed at random. Turnstile– elicited links compare well with survey–elicited links, albeit the large false–negatives rate. More importantly, the characteristics of turnstile–elicited links are statistically the same as those from the survey links, but different to those that would be obtained if links were obtained purely at random (see Figure 9). I rationalize measurement error in a difference– in–difference 2x2 framework that follows Goodmann–Bacon (2019) and Cunninghan (2021). If the measurement error is associated with the exposure to low–income students in ways unobserved by the researchers, then the observed Average Treatment on the Treated (ATT) effect may differ from the true ATT. I proxy measurement error with the number of ID taps on the turnstiles and the number of courses taken with turnstile–elicited links. I do

---

[13]The survey was conducted online between December 7, 2017, and January 5, 2018, and elicited the network among 110 economics students from the 2017 fall cohort. The survey was conducted using Qualtrics. Students who completed the survey received a free lunch voucher for a recognized chain restaurant of the campus area. Cárdenas et al. (2019) provide a detail description of the survey. I am very grateful to Professor Tomás Rodríguez-Barraquer for providing me access to these data.

not find evidence of the change in the percentage of low–income peers affecting any of these measures (see Table 9). In summary, I do not find evidence that measurement error in turnstile–elicited interactions bias my estimates on the impacts of exposure.

6   Results

Table 3 displays the estimated effect of the increased exposure to low–income peers on relatively wealthy students' academic achievement and persistence. Panel A displays OLS estimates of β_l for all the outcomes discussed in Figures 4 and 5. The size of the estimated effect on wealthy students' achievement is modest and not statistically different to zero for GPA and graduation outcomes. The point estimates for the number of credits attempted by first and third terms and for dropout probability by the 5th term are positive and statistically significant, but the size of the point estimate is small relative to the standard deviation of the variables. Namely, I estimate 0.02 additional credits attempted by the first term and 0.3 credits by the third term (0.04 standard deviations), and an increase in dropout probability 0.2 percentage points (0.06 standard deviations) when evaluated at the mean change $\Delta mu(R_{mc}^l)$=9.5 percentage points. Importantly, the estimated effect on dropout does not pair with impacts on graduation probability, which are estimated to be zero. These results suggest the increased exposure to low–income peers has if anything a positive yet small impact on the academic achievement of wealthy students, albeit a small negative effect on persistence which has not translated on impacts on graduation.

Prior literature has found non–linear effects of peers on academic achievement (Garlick, 2018; Zimmerman, 2003). Hence, Panel B of Table 3 estimates the effect of exposure when the increased percentage is at the top 25th percentile of the distribution (i.e., use a indicator equal to one when the percentage of low–income peers exceeds 30 percent). I do not find evidence of non–linear effects, which suggest even students in groups with the largest shares of peers are not affected in their academic achievement.

In summary, I do not find evidence that exposure to low–income peers has negative impacts on the academic achievement of wealthy students and if anything, I find suggestive evidence of a positive impact on the number of credits attempted, albeit a small yet significant impact on academic persistence as measured by dropout probability by the 5th term of college. Importantly, the measures of persistence are relatively premature, as by the time of the data collection only few students had started to graduated (for example, graduation in 8 terms –which is considered graduation on time, was only at the 7 percent among wealthy students). Plus, my measure of dropout considers only college–level dropout as I am unable to see whether these students are enrolled at a different university, meaning transfers to other universities cannot be ruled out. Overall, my findings speak to prior literature finding no effects of desegregation on receiving students achievement in K–12 settings(Angrist and Lang, 2004), and contradict some of the findings from higher education settings (Arcidiacono and Vigdor, 2010).

Exposure to low–achievers could impact the performance of wealthy students, even after accounting for the percentage of low–income students in the group. For example, low–achievers could also exhibit disruptive behaviors impacting wealthy students (Carrell et al., 2018), and they could reduce the incentives of wealthy students to exert effort. Figure 4 showed low–income students brought by SPP have a significantly lower academic achievement than their wealthy peers. Moreover, Table 1 showed incoming low–income students had on average a lower performance than wealthy students even before enrolling in college, as shown by their average SB11 score. To test this, I expand Equation 1 by including average measures of achievement of low–income peers, namely their average achievement outcome and their average SABER 11 test score. Specifically, I estimate:

$$Y_i^{mc} = \eta_{l2}R_{mc}^l + \eta_{S\tilde{B}11}S\tilde{B}11_{mc} + \mathbf{X_i'}H_2 + \eta_{2m} + \eta_{2c} + u_{2imc} \tag{2}$$

$$Y_i^{mc} = \eta_{l1}R_{mc}^l + \eta_{\tilde{Y}}\tilde{Y}_{mc} + \mathbf{X_i'}H_1 + \eta_{1m} + \eta_{1c} + u_{1imc} \tag{3}$$

20

As coined in the peer effects and Linear–in–Means–Models literature (Manski, 1993; Moffitt, 2001), Equation 2 captures the exogenous effect of low–income prior achievement on wealthy students performance, and Equation 3 captures the endogenous effect of low–income performance on wealthy students' performance –that is, the relation between the outcome of low–income peers with those of wealthy peers. Specifically, $S\tilde{B}11_{mc} = \frac{\sum_{j \neq i}^{N}[SB11_j|W_j=0]}{N_{mc}^l}$, and $\tilde{Y}_{mc} = \frac{\sum_{j \neq i}^{N}[Y_j|W_j=0]}{N_{mc}^l}$ where $W_j = 0$ if the peer j in student i major m and entry cohort c is low–income, and N_{mc}^l is the number of low–income students in student i major and entry cohort. Estimates of $\eta_{S\tilde{B}11}$ and $\eta_{\tilde{Y}}$ capture whether low–income students' SB11 test scores and academic achievement explain wealthy students' performance.

Table 4 displays the results of this analyses. Panel A displays results from estimating Equation 2. The estimates of $\eta_{S\tilde{B}11}$ are different from zero in the case of first term credits and for sixth term GPA, suggesting low–income peers' SB11 is positively associated with the number of credits attempted in the first term by wealthy students. However, the sign of the estimate inverts to negative in the sixth term. Panel B displays the results from 3. With the exception of first term credits attempted, the estimated $\eta_{\tilde{Y}}$ is not different to zero for achievement and persistence outcomes. Overall, these results suggest the performance of low–income students has some relation with the number of credits attempted in the first term by wealthy students, but the effect disappears by subsequent terms. Specifically, SB11 test scores are positively associated with more credits attempted in the first term by wealthy students, by more credits attempted in the first term by low–income students has a negative association with those of wealthy students.

The estimated effects of exposure on the academic achievement of wealthy students may be small because of a lack of interactions between wealthy and low–income students. Findings from Carrell et al. (2013) suggest assignment of low–achieving students to groups can lead to segregation among high–achievers, which may have unintended consequences for both groups. Thus, if the low achievement of incoming low–income students leads to segregation across income groups, the potential impacts on academic achievement may be zero, and the

effects other non–academic outcomes found by previous literature (Boisjoly et al., 2006; Rao, 2019) may be diminished or even absent. To test for this, I proceed to estimate the effects of exposure to low–income students on the diversity of social interactions of wealthy students.

Table 5 displays the results of estimating Equation 1 on wealthy students interactions with their peers. I measure three types of outcomes: the number of links of wealthy students with any of their peers, with other wealthy peers and with low–income peers in their major and entry cohort (Panel A); the probability of any link with wealthy or low–income peers (Panel B); and the proportion of links with low–income peers (Panel C). Panel A can be thought off as the impacts in the intensive margin i.e., the effect of exposure to low–income peers on the probability of an additional link, whereas panel B can be read as the impacts in the extensive margin i.e., the effects of exposure on the probability of forming any link with low–income peers. Panel C can be read as the effects on friendships compositions relative to the size of low–income population in the group. Similar to the effects on achievement and persistence, Panel 1 of the table shows linear estimates following Equation 1, whereas Panel 2 shows non–linear estimates that compute the effects when exposure to low–income peers is at the top 25th percent of the distribution in 2015 (i.e., use a indicator equal to one when the percentage of low–income peers exceeds 30 percent).

Exposure to low–income peers has a positive impact on the diversity of interactions of wealthy peers. Focusing on Panel 1 of Table 5, I find that the average increase in the percentage of low–income peers (9.51 percentage points) augmented the number of links with low–income peers in 0.28 and reduced the number of links with other wealthy peers in 0.41 on average. Relative to the 2014 cohorts, this represents an 0.51 standard deviation (s.d.) increase in the number of links with low–income peers and an 0.08 s.d. decrease in the number of links with other wealthy peers. Importantly, the increased exposure to low–income peers did not affect the total number of links of wealthy students. Moreover, the estimated effects persist at the extensive margin. That is, the average increase in the percentage of low–income peers increased the probability of having a low–income link in 0.7 percentage

22

points and decreased the probability of a link with another wealthy peer in 0.02 percentage points. Respectively, these are effects represent a 0.2 and 0.005 standard deviations change relative to the 2014 distribution. Lastly, I estimate the effect on the percentage of links with low–income students. A 9.5 percentage points increase in the percentage of low–income peers translates in a 6.8 percentage points increase in the percentage of links with the low–income. This is equivalent to a 0.6 s.d. effect relative to the outcomes of the 2014 cohorts.

The effects of exposure to low–income peers on the diversity of social interactions are significantly larger at the top 25th percentile of the distribution. Students who were exposed to shares of low–income peers over the 30 percent have 0.68 more low–income links, have an increase in the probability of having at least one low–income link of 1.2 percentage points, and have an increase in the percentage of low–income links of 14.57 points. Hence, the positive effects on diversity of social interactions persist even for those students experiencing the largest exposure to low–income peers.

In summary, exposure to low–income students have significant and positive impacts on the diversity of social interactions of wealthy students, as measured by their number, probability and percentage of links with the low–income. Relative to the number of links from the pre–treatment period the effects are estimated to be about half standard deviation for the number and percentage of links with the low–income and 0.2 standard deviations for the probability of a link with a low–income. Importantly, my results are robust to alternative definitions of turnstile–elicited interactions as shown in Appendix C. Overall, these results are larger than those found by previous approaches in the literature that have measured the intensity of interactions across different race groups (Marmaros and Sacerdote (2006) and Boisjoly et al. (2006) find positive impacts on the intensity of interactions ranging between 0.06 and 0.35 standard deviations), but with the advantage that my definitions of social interactions account for the probability and the number of different links according to their socio–economic background.

7 Analysis of the Results and Theoretical Predictions

The theoretical predictions discussed in Section 2 indicate that wealthy students would diversify their social interactions if those can be with low–income high–achieving students. This is because wealthy students would not anticipate adverse spill over effects from these students, hence, prior predictions from the literature suggesting low–achievement as a driver of segregation would not hold (Carrell et al., 2013). To test this prediction, I identify the low–income links of wealthy students whose performance is above the average of that of wealthy students in the major and entry cohort. My results suggest about half of the impact of exposure on the number of links with low–income peers is explained by links with high achievers. Plus, over 75 percent of the effect on the probability of interactions with low–incomes can be explained by the likelihood of interactions with high achieving low–income students.

Table 6 displays the results from this analysis. Similar to Table 5, I compute the effects of the changes in exposure to low–income students on the number and probability of a link with a low–income student who is also a high achiever in terms of: SB11 test scores, first term GPA, and total credits attempted in the first term. I consider a low–income student to be high achiever if their performance in the achievement variable is above the average of that of the wealthy students in their major and entry cohort. Theoretically, wealthy students would observe the performance of their low–income peers and compared it to that of their wealthy classmates in their major and cohort. If that performance is above average, then the wealthy student would anticipate no harm to their achievement or, if anything, a positive peer effect. Hence, favoring interactions with the low–income.

Focusing on Panel A of Table 6, I find that the average increase in exposure to low–income students of 9.5 percentage points led to 0.12 more links with low–income students with above average SB11 test scores, 0.15 more links with low–income students with above average first term GPA, and 0.20 more links with low–income students with above average credits attempted in the first term. Relative to the initial estimated effect of exposure on

links with any low–income student of 0.28, the different measures of achievement explain 41 percent, 52 percent and 68 percent of the total effect, respectively. Importantly, the fact that 68 percent of the links with the low–income can be explained by the number of credits attempted maps well with the positive but modest impacts of exposure on credits attempted by wealthy students in their first and third terms of college. Notably, these results are suggestive of positive effects of diversifying interactions on the achievement of wealthy students.

Panel B of Table 6 displays the impacts on the probability of interaction with the low–income by the different achievement measures. Wealthy students exposed to an increase of low–income peers of 9.5 percentage points are 0.6 percentage points more likely to interact with a low–income student with above average SB11 achievement, and 0.6 percentage points and 0.76 percentage points more likely to interact with a low–income student with above average first term GPA and credits attempted, respectively. Relative to the initial estimated impact of exposure on the likelihood of interaction with any low–income peer, I find that high achievement in terms of SB11 and GPA among low–income students explains 75 percent of the effect on the probability of interaction with any low–income student, and almost a 100 percent of the probability when achievement is measured in term of credits attempted in the first term. Coupled with the results on achievement, matching in terms of the number of credits attempted in the first term is a key driver of the diversity in social interactions driven by exposure.

8 Conclusions

In 2014, the Colombian government launched *Ser Pilo Paga* a financial aid program targeting low–income high–achieving students that provided a forgivable loan covering tuition plus a stipend for living expenses, so long as the eligible students attended high-quality selective universities in the country. The program triggered a sharp influx in the number

of low–income students enrolling at elite institutions in 2015. I study the case of one Elite University in Colombia where the enrollment of low–income students tripled with SPP, but that maintains the number of wealthy students enrolling in the 2015 cohort. The policy induced a quasi–random variation in the percentage of low–income students at each major and entry cohort, which I exploit in a difference–in–differences design to answer: can exposure to desegregation at an elite college lead to more diverse social interactions? Could it do so without harming the achievement of students traditionally attending the institution?

I do not find evidence suggesting exposure to low–income students through desegregation policies have adverse impacts on the achievement of students at the receiving institutions, thus debating previous results by Arcidiacono and Vigdor (2010) and providing empirical evidence debating the predictions from the peer effects literature (Arcidiacono et al., 2015). Moreover, I find significantly positive yet minor impacts on early achievement as measured by the cumulative credits attempted by the first and third college terms. Importantly, I also find significantly positive impacts on dropout probability by the fifth term of college. However, my dropout measure is captured at the college level, and my data does not allow me to track these students at other institutions. Thus, transfer to other universities instead of total drop out of higher education cannot be ruled out.

Given that low–income students enrolling at the outset of SPP have significantly lower academic achievement relative to their wealthy peers, I also inspect for peer effects driven by exposure to low achievement, finding no evidence in this regard. Overall, my findings are aligned with those from Angrist and Lang (2004), and Dobbie and Fryer (2014) in that I do not find significant adverse impacts of desegregation on achievement. Importantly, my findings also contribute to the peer effects literature in higher education, which so far has found mixed results of achievement peer effects (Zimmerman, 2003; Garlick, 2018).

Lack of impacts on academic achievement from exposure to desegregation may be explained by lack of social interactions among wealthy and low–income students. Such behavior would be consistent with prior research suggesting changes in peer effects composition

driven by incoming students with lower achievement may induce segregation within groups (Carrell et al., 2013). Thus, I examine whether exposure to desegregation does induce more interactions between wealthy and low–income students. I find that the average increase in the percentage of low–income students induced significant socioeconomic desegregation in social interactions. The number of links between wealthy and low–income students increased by 0.28 (0.51 standard deviations relative to the pre–SPP entry cohort), and the overall probability of interaction between a wealthy and a low–income student increased by 0.7 percentage points (0.2 standard deviations relative to the pre–SPP entry cohort). Overall, this implies a 6.8 percentage points increase in the percentage of links that wealthy students have with low–income peers.

These findings contribute to the literature that examines the role of group membership in college in explaining employment and social mobility outcomes (Zimmerman, 2019; Michelman et al., 2020; Marmaros and Sacerdote, 2002). Specifically, while membership to groups in college can be a proxy of social interactions, it is also confounded by other factors such as signaling value to the labor market. In that sense, this paper provides evidence on how social interactions work at elite colleges, using a robust measure of students interactions.

Lastly, I test theoretical predictions explaining that exposure increases the diversity of social interactions without harming academic achievement. Wealthy students may assess the achievement of low–income students relative to that of their wealthy peers to establish a link. If students segregate to avoid adverse peer effects (Carrell et al., 2013), then students integrate if they do not perceive adverse impacts on their achievement. I find suggestive evidence supporting this hypothesis. At least half of the links with low-income students are with students whose average achievement is above that of the wealthy students in the group. Importantly, 68 of the effect on the number of links with low–income students is explained by links with high achievers in terms of attempted credits. These findings suggest that desegregation in social interactions is strongly driven by academic achievement and that low–income high–achieving students have a large advantage in achieving integration relative

to other low-income students.

The results in this paper provide substantial evidence in favor of desegregation policies at the higher education level. Exposure to desegregated environments can lead to diversity in interactions without adverse effects on achievement. Desegregation in social interactions is important as it can nurture access to information and social and cultural capital, which can be determinant in the job market (Rivera, 2016). Future research should inquire whether the effects of desegregation can translate with impacts in the quality of employment of the low–income students benefiting from desegregation.

References

Angrist, J. D. and Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's metco program. *American Economic Review*, 94(5):1613–1634.

Angrist, J. D. and Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press, Princeton, NJ., first edit edition.

Arcidiacono, P., Lovenheim, M., and Zhu, M. (2015). Affirmative Action in Undergraduate Education. *Annual Review of Economics*, 7(1):487–518.

Arcidiacono, P. and Vigdor, J. L. (2010). Does the river spill over? estimating the economic returns to attending a racially diverse college. *Economic Inquiry*, 48(3):537–557.

Baker, S., Mayer, A., and Puller, S. L. (2011). Do more diverse environments increase the diversity of subsequent interaction? Evidence from random dorm assignment. *Economics Letters*, 110(2):110–112.

Billings, S. B., Chyn, E., and Haggag, K. (2021). The Long-Run Effects of School Racial Diversity on Political Identity. *American Economics Review: Insights*, 3(3):No. 27302.

Bleemer, Z. (2020). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or antipathy? The impact of diversity. *American Economic Review*, 96(5):1890–1905.

Camacho, A., Messina, J., and Uribe, J. P. (2017). The Expansion of Higher Education in Colombia: Bad Students or Bad Programs?

Canaan, S. and Mouganie, P. (2018). Returns to Education Quality for Low-Skilled Students: Evidence from a Discontinuity. *Journal of Labor Economics*, 36(2).

Cárdenas, J. C., Okulicz, D., Pietrobon, D., and Rodríguez, T. (2019). Propensity to Trust and Network Formation.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415.

Carrell, S. E., Sacerdote, B., and West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81(3):855–882.

Castleman, B. L. and Long, B. T. (2016). Looking beyond Enrollment: The Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income Segregation and Intergenerational Mobility Across Collegres in the United States. *Quarterly Journal of Economics*, pages 1567–1633.

Christakis, N., Fowler, J., Imbens, G., and Kalyanaraman, K. (2010). An Empirical Model for Strategic Network Formation. *National Bureau of Economic Research*.

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

Dale, S. B. and Krueger, A. B. (2002). Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *The Quarterly Journal of Economics*, 117(4):1491–1527.

Dobbie, W. and Fryer, R. G. (2014). The impact of attending a school with high-achieving peers: Evidence from the New York City exam schools. *American Economic Journal: Applied Economics*, 6(3):58–75.

Echenique, F., Fryer, R. G., and Kaufman, A. (2006). Is school segregation good or bad? *American Economic Review*, 96(2):265–269.

Fack, G. and Grenet, J. (2015). Improving college access and success for low-income students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics*, 7(2):1–34.

Garlick, R. (2018). Academic peer effects with different group assignment policies: Residential tracking versus random assignment. *American Economic Journal: Applied Economics*, 10(3):345–369.

Hoekstra, M. (2009). The Effect of Attending the Flagship State University on Earnings: a Discontinuity-based Approach. *The review of Economics and Statistics*, 91(4):717–724.

Londoño-Vélez, J. (2020). The Impact of Diversity on Distributive Perceptions and Preferences for Redistribution.

Londoño-Velez, J., Rodriguez, C., and Sanchez, F. (2020). Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: Ser Pilo Paga in Colombia. *American Economic Journal: Economic Policy*, 12(2):193–227.

Long, M. C. (2008). College quality and early adult outcomes. *Economics of Education Review*, 27(5):588–602.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. Technical Report 3.

Marmaros, D. and Sacerdote, B. (2002). Peer and social networks in job search. *European Economic Review*, 46:870–879.

Marmaros, D. and Sacerdote, B. (2006). How Do Friendships Form? *The Quarterly Journal of Economics*, 121(1):79–119.

Marta Ferreyra, M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., and Urzúa, S. (2017). *At a Crossroads Higher Education in Latin America and the Caribbean Human Development*. The Worl Bank, Washington D.C.

Mele, A. (2017). A Structural Model of Dense Network Formation. *Econometrica*, 85(3):825–850.

Mele, A. (2020). Does school desegregation promote diverse interactions ? An equilibrium model of segregation within schools. *American Economic Journal: Economic Policy*, 12(2).

Mello, U. (2021). Centralized Admissions, Affirmative Action and Access of Low-income Students to Higher Education. *American Economic Journal: Economic Policy*, forthcomin.

Michelman, V., Price, J., and Zimmerman, S. D. (2020). The Distribution of and Returns to Social Success at Elite Universities.

Moffitt, R. A. (2001). Policy Interventions, Low-Level Equilibria, and Social Interactions. In Durlauf, S. and Young, P., editors, *Social Dynamics*. MIT Press.

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

Rogers, B. W. and Jackson, M. O. (2007). Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review*, 97(3):890–915.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1):9–23.

Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review*, 109(1):1–47.

Tables

Table 1: Descriptive Statistics of Students' Characteristics

	2014 entry cohort				2015 entry cohort			
	Wealthy	Low–income			Wealthy	Low–income		
	Mean	Mean	Difference		Mean	Mean	Difference	
Peers composition								
Prop. Of middle–SES	0.49	0.00	-0.49	***	0.50	0.00	-0.50	***
No. Of links	5.21	4.94	-0.27		5.53	4.60	-0.93	**
No. of low SES links	0.24	0.35	0.11	*	0.59	1.73	1.14	***
Student characteristics								
Female	0.43	0.35	-0.09	**	0.46	0.41	-0.06	
Age	17.59	17.24	-0.35	***	17.59	17.14	-0.45	***
Mother with no college degree	0.08	0.24	0.16	***	0.09	0.40	0.32	***
SB11 standardized test score	0.00	-0.09	-0.10		0.04	-0.22	-0.26	***
SPP recipient	0.00	0.00	0.00		0.05	0.81	0.76	***
Other scholarship recipient	0.07	0.37	0.30	***	0.07	0.05	-0.02	
Immigrant according to H.S.	0.23	0.35	0.12	**	0.23	0.56	0.33	***
No. Of peers from H.S. in same cohort	11.54	3.17	-8.38	***	11.73	1.98	-9.75	***
Turnstiles characteristics								
Tot. ID taps in turnstiles in 6th and 7th terms	1340.19	1349.79	9.60		1239.93	1162.33	-77.61	
Links' characteristics								
Age Difference	0.60	0.65	0.06		0.63	0.63	0.00	
Same Gender	0.50	0.51	0.00		0.51	0.45	-0.05	**
Ave. No. Of courses w/links	1.49	1.37	-0.13		1.51	1.31	-0.20	
SB11 difference	0.73	0.76	0.03		0.79	0.67	-0.12	**
Share of friends from same high school	0.04	0.01	-0.03	***	0.04	0.00	-0.04	***
No. Of majors	31	31			31	31		
No. Of students	2669	139			2609	538		
Individuals without links with their major-cohort	600	40			541	151		

Notes: This table displays descriptive statistics of the sample of students described in Section 4. Wealthy students comprise middle– and high–SES students. * p < 0.10, ** p < 0.05, *** p < 0.01. P-values are based on t–test of the hyphotesis that the difference in means between wealthy and low–income students is equal to zero. Standard errors where clustered at the major level.

Table 2: Identification Challenges in the Difference–in–Difference Research Design

Type of Bias	Source	Assessment	Conclusion
Non–parallel trends bias	**Selection of majors and entry cohorts changes with SPP**	Event study on outcomes and co-variates (see Figures 7, 8, and 6)	No evidence of bias.
	Wealthy students are crowed out in the admission process by the low-income students recipients of SPP.	Total number of seats to high-income students did not change (see Figure 2)	There is no correlation between the increased number of low–income students per program and cohort and the change in the number of wealthy students
Unobserved exposure effects	**Exposure to low–SES operates through courses, and many courses can be taken by students from multiple majors.**	Index capturing exposure to low–SES peers at the course level. I embed the index in a within major-cohort specification and use IV to ensure exogeneity (see Appendix B).	Course-level exposure does not distinctly affect outcomes for students within the same major and cohort
Measurement Error on the social interactions outcomes	**Turnstile-elicited interactions are the result of random co–movements**	Use data from a survey on social interactions among students from one major and entry cohort (see Appendix A - Tables 7 and 8).	The rate of false positives (i.e., turnstile–elicited interactions not captured by the survey) is below 10 percent
	If false-negatives are non-random, there is bias in the turnstile-elicited links	I use simulations to estimate how far from random are the turnstile-elicited links' characteristics, and how close the average characteristics of the turnstile–elicited links are to those of the survey-elicited links (see Appendix A - Figure 9).	Turnstile–elicited interactions capture links characteristics, albeit the number of interactions missing
	Rate of false positives and negatives across majors and cohorts could be impacted by exposure to low–income peers	Use proxies of measurement error to test for the effects of exposure using the Dif–in–Dif design (See Appendix A - Table 9).	I do not find evidence of an effect of exposure on measurement error rates.

Notes: This table describes the identification challenges to the Difference–in–Difference research design discussed in Equation 1.

Table 3: The Impact of Exposure to Desegregation on Academic Achievement and Persistence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st term credits	1st term GPA	3rd term cum. Credits	3rd term cum. GPA	6th term cum. Credits	6th term cum. GPA	Dropout by 5th term	Graduated in 8 terms
A. OLS								
R^l_{mc}: % of low–income peers	0.011*	0.001	0.034**	0.001	0.031	0.001	0.002**	-0.000
	(0.006)	(0.001)	(0.017)	(0.001)	(0.031)	(0.001)	(0.001)	(0.001)
B. Non-linear Effects								
$\mathbb{I}$[% of low–income peers > 30%]	0.063	-0.021	0.277	0.002	-0.065	-0.004	0.028	-0.034
	(0.174)	(0.033)	(0.571)	(0.026)	(0.815)	(0.022)	(0.020)	(0.029)
mu(Y)	15.64	3.863	49.39	3.822	100.9	3.859	0.122	0.0693
sd(Y)	2.949	0.449	8.496	0.378	16.22	0.344	0.327	0.254
Treatment distribution in 2015								
Δ mu(R^l_{mc})				9.51 p. p.				
sd(R^l_{mc})				12.98				
No. Students	5,278	5,274	4,895	4,895	4,507	4,507	5,278	5,278

Notes: This table displays the estimates from Equation 1 in Panel A., and the non–linear estimates in Panel B. *% of low–income peers* is calculated at the major–cohort levels. $\mathbb{I}$[% of low–income peers > 30%] is an indicator function equal to one if the share of low–income peers in the major and cohort is greater than 30% i.e., over the 75th percentile of the distribution in the 2015 – Spring entry cohort. All estimations control for female indicators, age in year at the time of entry, SB11 standardized test score, mothers' has no college education indicator, indicator of middle–SES background according to the social strata indicator, and dummies for whether the student got an SPP loan. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 4: Endogenous and exogenous effects of low–income peers on wealthy students' achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st term credits	1st term GPA	3rd term cum. Credits	3rd term cum. GPA	6th term cum. Credits	6th term cum. GPA	Dropout by 5th term	Graduated in 8 term
A. Exogenous SB11								
R^l_{mc}	0.014**	0.001	0.034**	0.001	0.031	0.000	0.002**	-0.000
	(0.005)	(0.001)	(0.017)	(0.001)	(0.032)	(0.001)	(0.001)	(0.001)
$S\tilde{B}11_{mc}$	1.262***	-0.006	0.204	-0.085	-0.058	-0.119**	-0.018	0.019
	(0.439)	(0.052)	(1.027)	(0.057)	(2.415)	(0.053)	(0.046)	(0.044)
B. Endogenous Achievement								
R^l_{mc}	0.013**	0.001	0.033*	0.001	0.030	0.001	0.002**	-0.000
	(0.006)	(0.001)	(0.017)	(0.001)	(0.032)	(0.001)	(0.001)	(0.001)
$\tilde{Y}_{mc}$	-0.015*	-0.003	0.004	0.005	0.003	0.006	0.032	0.021
	(0.009)	(0.005)	(0.007)	(0.005)	(0.008)	(0.004)	(0.045)	(0.028)
mu(Y)	15.64	3.866	49.43	3.825	100.9	3.862	0.121	0.0692
sd(Y)	2.939	0.448	8.469	0.377	16.19	0.344	0.326	0.254
No. Students	5,278	5,274	4,895	4,895	4,507	4,507	5,278	5,278

Notes: This table displays the estimates from Endogenous Peer Effects (Eq. 3) and Exogenous SB11 Effects (Eq: 2). N_{Pmc} is the proportion of low–income students in the major and entry cohort as treatment. $S\tilde{B}11_{mc} = \frac{\sum_{j \neq i}^{N}[SB11_j | W_j = 0]}{N^l_{mc}}$, and $\tilde{Y}_{mc} = \frac{\sum_{j \neq i}^{N}[Y_j | W_j = 0]}{N^l_{mc}}$ where $W_j = 0$ if the peer j in student i major m and entry cohort c is low–income, and N^l_{mc} is the number of low–income students in student i major and entry cohort. All regressions include the controls described in Equation 1. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 5: The Impact of Exposure to Desegregation on The Links of Wealthy Students

	A. Number of Links with:			B. Probability of a Link with:		C. Proportion of Links with:
	(1) Any	(2) Wealthy	(3) Low Income	(4) Wealthy	(5) Low Income	(6) Low Income
1. OLS						
R^l_{mc}: *% of low–income peers*	-0.011	-0.043***	0.031***	-0.002**	0.008***	0.714***
	(0.016)	(0.016)	(0.004)	(0.001)	(0.001)	(0.060)
2. Non-linear Effects						
$\mathbb{I}$[% of low–income peers > 30%]	0.017	-0.667	0.684***	-0.043	0.117***	0.146***
	(0.495)	(0.507)	(0.131)	(0.028)	(0.038)	(0.029)
Pre-treatment statistics of the outcomes						
mean	5.212	4.973	0.239	0.770	0.188	4.404
standard deviation	5.154	4.922	0.558	0.421	0.391	11.34
Treatment distribution in 2015						
Δ mu(R^l_{mc})				9.51 p. p.		
sd(R^l_{mc})				12.98		
No. Students	5,278	5,278	5,278	5,278	5,278	4,137

Notes: This table displays the estimates from Equation 1 in Panel A., and the non–linear estimates in Panel B. Outcomes are based on a turnstile–elicited links based on time–windows of 3 seconds and at least two co–movements in a term. *% of low–income peers* is calculated at the major–cohort levels. $\mathbb{I}$[% of low–income peers > 30%] is an indicator function equal to one if the share of low–income peers in the major and cohort is greater than 30% i.e., over the 75th percentile of the distribution in the 2015 – Spring entry cohort. All estimations control for female indicators, age in year at the time of entry, SB11 standardized test score, mothers' has no college education indicator, indicator of middle–SES background according to the social strata indicator, and dummies for whether the student got an SPP loan. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.
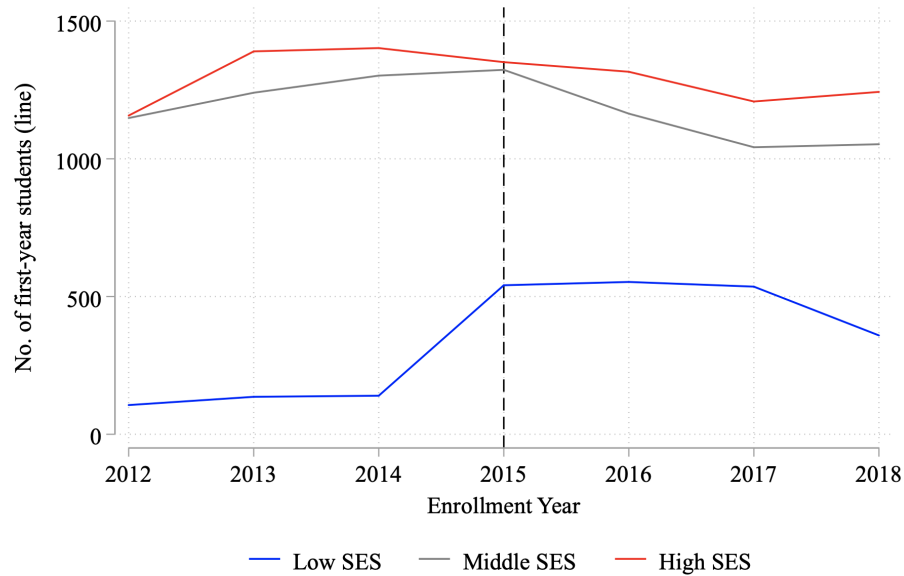
Table 6: The Impact of Exposure on Links with the High–Performing Low–Income Students

	A. Number of Links with High Performers By:			B. Probability of a Link with High Performers By:		
	(1) SB11	(2) First Term GPA	(3) First Term Credits	(4) SB11	(5) First Term GPA	(6) First Term Credits
OLS						
% of low–income peers	0.013***	0.016***	0.021***	0.006***	0.006***	0.008***
	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Pre-treatment statistics of the outcomes						
mean	0.0892	0.135	0.118	0.0798	0.111	0.102
standard Deviation	0.317	0.419	0.375	0.271	0.315	0.303
Treatment distribution in 2015						
Delta mu(R)			9.51 p. p.			
sd(R)			12.98			
No. Students	5,278	5,278	5,278	5,278	5,278	5,278

Notes: This table displays the estimates from Equation 1. Low–Income students are labeled as high achievers, when their performance is above the average of that of the wealthy students in their major and entry cohort. *% of low–income peers* is calculated at the major–cohort levels. All estimations control for female indicators, age in year at the time of entry, SB11 standardized test score, mothers' has no college education indicator, indicator of middle–SES background according to the social strata indicator, and dummies for whether the student got an SPP loan. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.
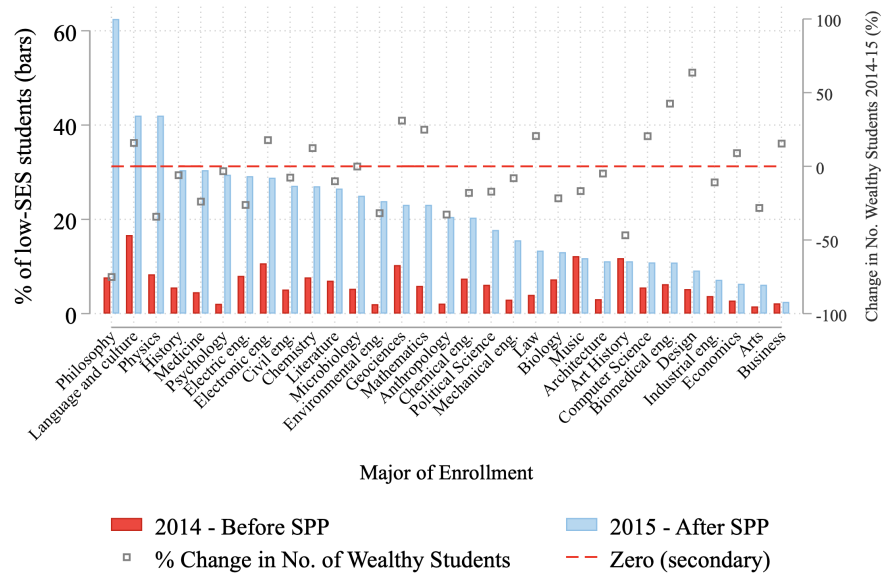
Figures

Figure 1: Number of first term students by socio–economic status (SES)

Figure 2: Percentage of low–SES students by major and before and after SPP

Figure 3: Flow of students by entrance - term and hour according to turnstiles

Hourly flow by entrance
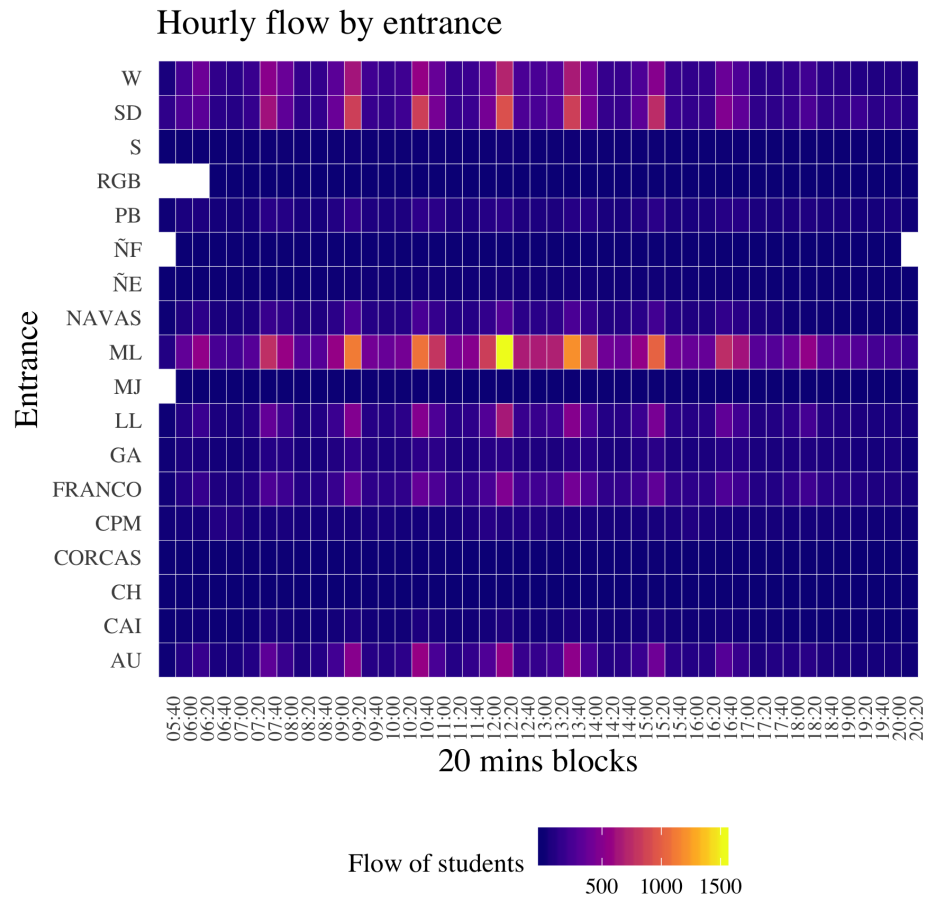
Notes: Average number of taps per day, entrance and 20 minutes blocks. Taps include INs and OUT of the building. Each term has approximately 75 weekdays

Figure 4: Average Cumulative GPA and Credits Attempted by Entry Cohort and Income group

First term GPA

First term credits

Cum. GPA by 3rd term

Cum. credits by 3rd term

Cum. GPA by 6th term

Cum. credits by 6th term

Low-Income — Wealthy
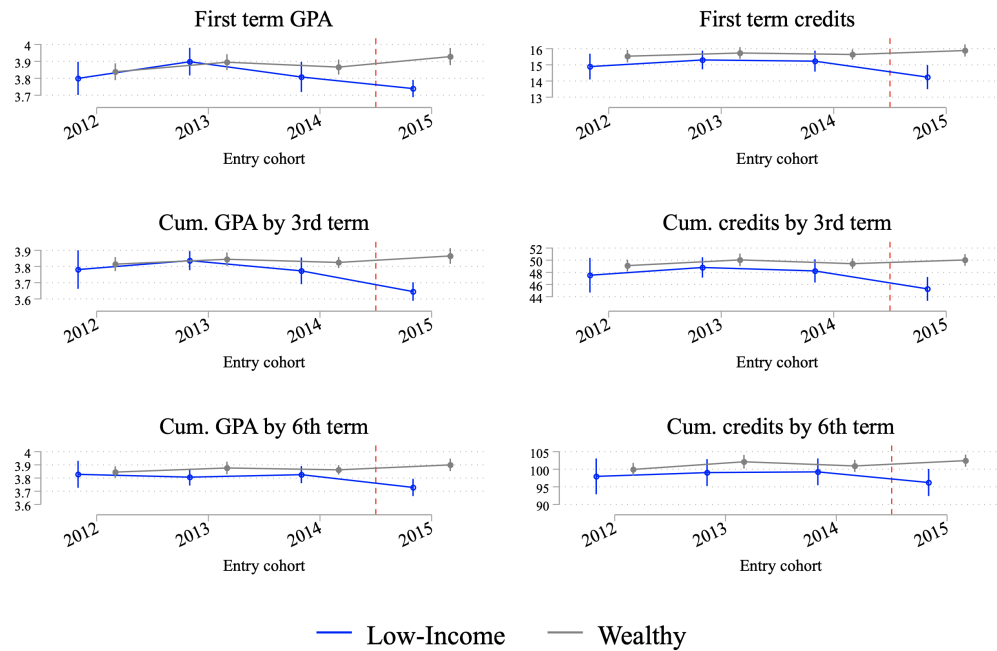
Notes: 95% confidence intervals based on clustered standard error by major.

Figure 5: Academic Persistence by Entry Cohort and Income group

Dropped Out by 5th term

Graduated in 8 term or less

Low-Income — Wealthy
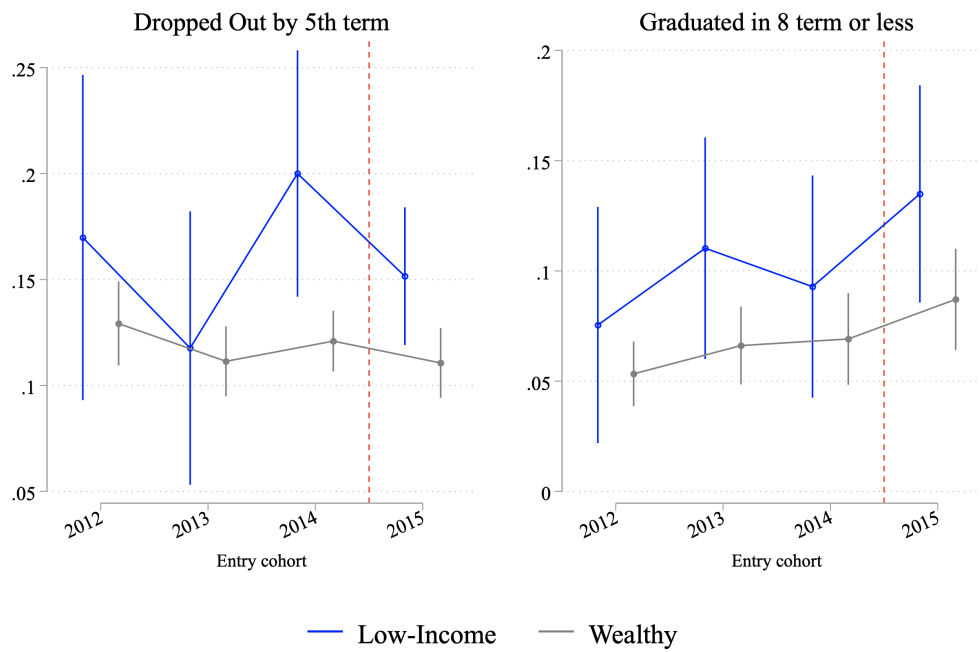
Notes: 95% confidence intervals based on clustered standard error by major.

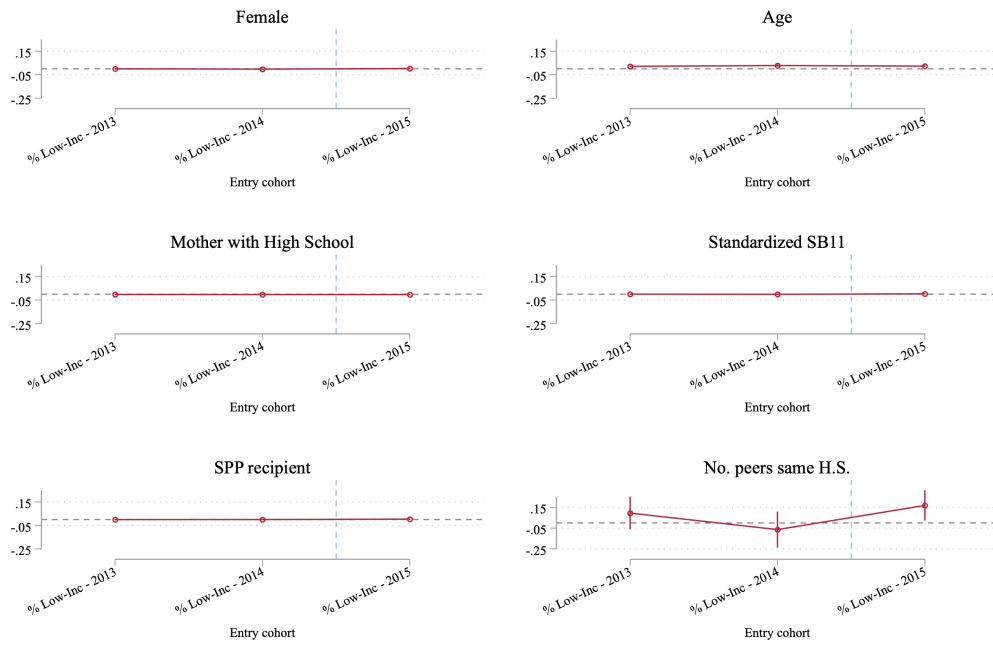Figure 6: Variation in Students' Characteristics within Cohorts and Across majors

43

Figure 7: Variation in Students' Achievement within Cohorts and Across Majors
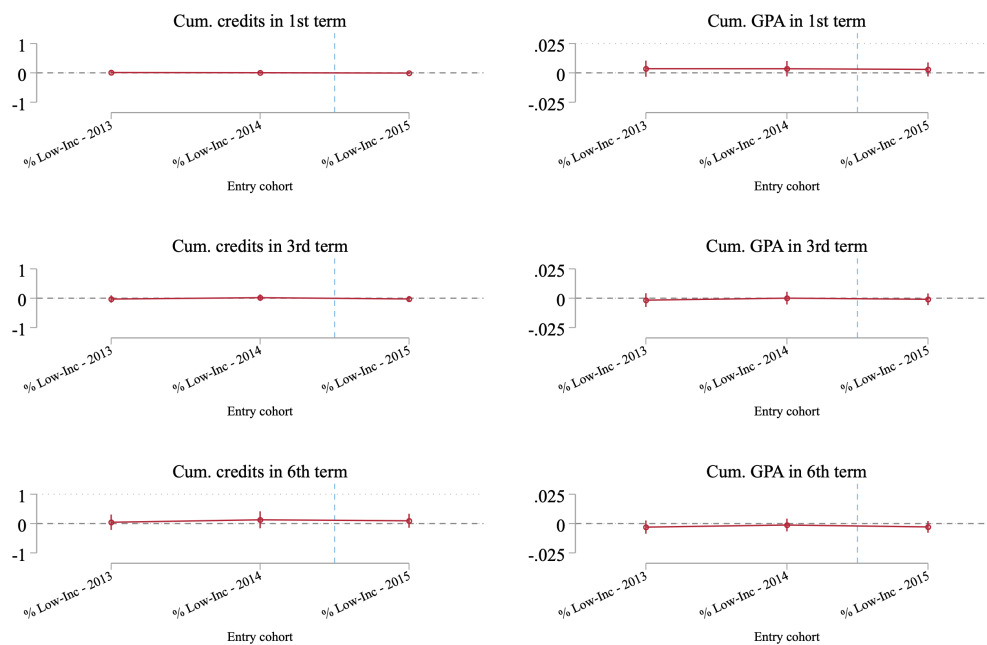
Notes: Point estimates of a cohort dummy from a regression where the dependent variable is the characteristic of the students controlling for major of enrollment dummies. Each year entry cohort includes the entry cohorts of Spring and Fall of the respective calendar year. Cluster standard errors at the major-cohort level. 95% confidence intervals.

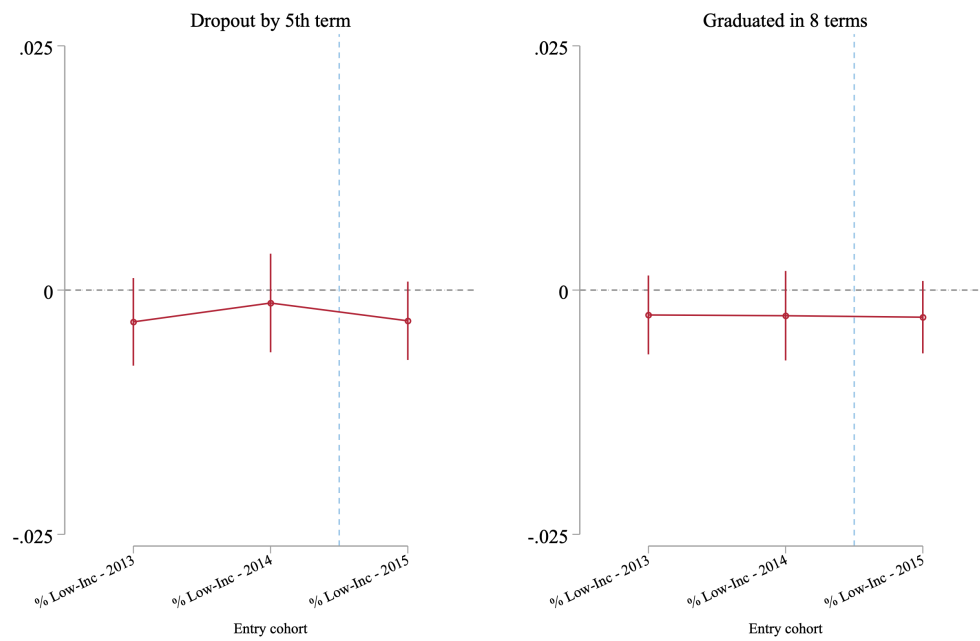Figure 8: Variation in Students' Persistence within Cohorts and Across Majors

Notes: Point estimates of a cohort dummy from a regression where the dependent variable is the characteristic of the students controlling for major of enrollment dummies. Each year entry cohort includes the entry cohorts of Spring and Fall of the respective calendar year. Cluster standard errors at the major-cohort level. 95% confidence intervals.

A   Appendix: Turnstile–Elicited Interactions Data and Validation

Validation of student links definition. I define a time window and frequency thresholds by comparing turnstile-elicited with survey-elicited links among first-term undergraduate students of Economics from the fall of 2017 cohort. The survey was conducted online between December 7, 2017, and January 5, 2018, and elicited the network among 110 economics students from the 2017 fall cohort. The survey was conducted using Qualtrics. Students who completed the survey received a free lunch voucher for a recognized chain restaurant of the campus area. Cárdenas et al. (2019) provide a detail description of the survey.[14]. The survey inquired about two types of links: friendships and acquaintances. Table 7 shows the results of the comparison. The time windows tested in Table 7 were selected based on in-person observations to different entrances. The observations of entrances to campus were conducted between August 26th and 30th of 2019. Because there are multiple turnstiles at each entrance, students walking together can essentially tap their IDs simultaneously using different scanners, thus the short time–windows. I select a time–window and a frequency criterion by minimizing the sum of the type II and type I measurement errors; that is, the number of unmatched survey-elicited links over the total number of survey links, and the number of unmatched turnstile-elicited links over the total number of survey links. For the purposes of this test, I assume the true number of links to which the type I and II errors refer are those captured by the survey.

To illustrate how to interpret the results in Table 7, I ask the reader to focus on the time window of three seconds and the acquaintances survey links. The numbers in bold indicate the combinations of time-windows and frequencies that minimize the sum of type I and II errors, for each type of link. Thus, the frequency with which I should observe two student IDs tapped on a turnstile entrance so that it resembles an acquaintances link should be minimum twice in the semester. Under that rule, the likelihood of Type I error or false positives - i.e. the likelihood of defining a pair of students as linked when according to the survey they are not, is 11 percent. Conversely, the likelihood of a Type II error or false negative - i.e., the likelihood of not identifying a pair of students as acquaintances when according to the survey they are, is 62 percent. While a five-seconds and three times in the term criteria would yield a lower sum of errors, it would do so by leaving one student from the 110 in the sample without turnstile-based links information –an omission I want to avoid. Notice that the acquaintances criteria has a lower threshold in terms of the frequency of the co-movements in the semester than the friendship criteria. I chose to use the acquaintances

---

[14]I am very grateful to Professor Tomás Rodríguez-Barraquer for providing access to these data.

instead of the friendship criteria because it allows me to identify social interactions that students did not identify as friendships in the first term of college, but that may eventually evolve as such.

The results in Table 7 indicate that under the baseline definition, it is highly likely that the turnstiles-elicited links capture survey-like links. However, an important share of survey links may not be captured by the turnstiles. This is an issue to the extent that those I do capture are not representative of the survey-elicited links. To assess this, I compare whether turnstile-elicited links plausibly reflect survey-elicited network characteristics. Results are displayed in Figure 9. The goal of this exercise is to estimate how far from random are the turnstile-elicited links' characteristics, and how close the average characteristics of the links are to those of the survey-elicited links. The computation proceeds as follows: I use the acquaintances minimizing criteria from Table 7 for each of the time windows and randomly assign the number of turnstile-elicited links under that criteria to the 110 students in the sample. Then, I compute the average of the following network individual attributes: age difference, number of courses students are taking together, GPA difference, degree or number of links, and local clustering. I conduct this procedure 1000 times and plot the distribution of the characteristics. I include the average value I observe for the turnstile- and survey-elicited links with its 95 percent confidence interval. I find statistically significant support indicating turnstile-elicited network characteristics resemble closely those of the friendship and acquaintances networks elicited by the survey, and are not the result of random links formation.

The validity of the turnstiles–elicited interactions could be susceptible to the hours of the day during which co–movements are captured. Co–movements captured around lunch hours could be more susceptible to false negatives, whereas co–movements captured at other times may be less susceptible to false positives. I test the extent to which this is an issue by replicating the comparison with the survey–elicited interactions from Table 7 but for co–movements happening around lunch–time hours (from 11:40 am to 2:20 pm) with co–movements at other times. The results are displayed in Table 8. For simplicity, I focus on Acquaintances links and on the two and three–seconds windows. As expected, co–movements captured during lunch–time are more susceptible to false negatives than co–movements captures off lunch–time. Similarly, co–movements captured at lunch–time are less susceptible to false positives than those captured at other times. However, the sum of error rates is much higher at either times than that obtained when using all times pooled together as presented in Table 7. These results suggest searching for co–movements at any time of the day is more reliable in terms of reducing measurement error than to focus on co–movements happening at specific times of the day.

Table 7: Survey– and Turnstile–elicited links comparison

Time window	A. two seconds					B. three seconds					C. Five seconds				
Frequency	One	Two	Three	Four	Five	One	Two	Three	Four	Five	One	Two	Three	Four	Five
1. Turnstiles															
No. Of dyads	868	368	235	180	148	1209	509	314	251	198	1906	898	552	401	315
No. of students	110	110	108	107	105	110	110	109	108	107	110	110	109	108	108
2. Are friends															
Dyads			*505*					*505*					*505*		
Survey & Turnstiles															
Matched	342	256	201	165	140	389	305	248	215	179	433	368	337	295	263
False Negatives (Type II)	0.32	0.49	0.60	0.67	0.72	0.23	0.40	0.51	0.57	0.65	0.14	0.27	0.33	0.42	0.48
False Positives (Type I)	1.04	0.22	0.07	0.03	0.02	1.62	0.40	0.13	0.07	0.04	2.92	1.05	0.43	0.21	0.10
Sum	1.36	0.71	**0.67**	0.70	0.74	1.85	0.80	**0.64**	0.65	0.68	3.06	1.32	0.76	0.63	**0.58**
3. Acquaintances															
Dyads			*1033*					*1033*					*1033*		
Survey & Turnstiles															
Matched	497	311	219	174	144	606	391	284	235	191	734	537	425	348	293
False Negatives (Type II)	0.52	0.70	0.79	0.83	0.86	0.41	0.62	0.73	0.77	0.82	0.29	0.48	0.59	0.66	0.72
False Positives (Type I)	0.36	0.06	0.02	0.01	0.00	0.58	0.11	0.03	0.02	0.01	1.13	0.35	0.12	0.05	0.02
Sum	*0.88*	***0.75***	*0.80*	*0.84*	*0.86*	*1.00*	***0.74***	*0.75*	*0.79*	*0.82*	*1.42*	*0.83*	***0.71***	*0.71*	*0.74*

Notes: N students = 110. Number of links possible (N*(N-1))/2 = 5995. Survey sample consist of economics undergrads from the August 2017 cohort. 113 students surveyed. One student did not report information and two do not show enrolled as of 2017-2. The survey asked each student who among the 113 students were an Acquaintance, and among those, who was considered a friend. Type II error rate is the share of links in survey that were not found in turnstiles-based links. Type I error is the links in turnstiles that were not matched with the links in survey, over the total links in survey

Table 8: Survey– and Turnstile–Elicited Links Comparison During and Off Lunch Time

| Time window	A. Two seconds						B. Three seconds					
Type	11:40 am to 2:20 pm			Other times			11:40 am to 2:20 pm			Other times		
Frequency	One	Two	Three	One	Two	Three	One	Two	Three	One	Two	Three
1. Turnstiles												
No. Of dyads	397	159	100	654	272	172	554	213	135	893	376	233
No. of students	110	109	103	110	110	105	110	109	106	110	110	106
2. Acquaintances												
Dyads			1033						1033			
Survey & Turnstiles												
Matched	255	143	93	411	236	162	321	180	123	494	308	214
False Negatives (Type II)	0.75	0.86	0.91	0.60	0.77	0.84	0.69	0.83	0.88	0.52	0.70	0.79
False Positives (Type I)	0.14	0.02	0.01	0.24	0.03	0.01	0.23	0.03	0.01	0.39	0.07	0.02
Sum	*0.89*	***0.88***	*0.92*	*0.84*	***0.81***	*0.85*	*0.91*	***0.86***	*0.89*	*0.91*	***0.77***	*0.81*

Notes: N students = 110. Number of links possible (N*(N-1))/2 = 5995. Survey sample consist of economics undergrads from the August 2017 cohort. 113 students surveyed. One student did not report information and two do not show enrolled as of 2017-2. The survey asked each student who among the 113 students were an Acquaintance, and among those, who was considered a friend. Type II error rate is the share of links in survey that were not found in turnstiles-based links. Type I error is the links in turnstiles that were not matched with the links in survey, over the total links in survey.

Figure 9: Comparison against randomly generated distribution
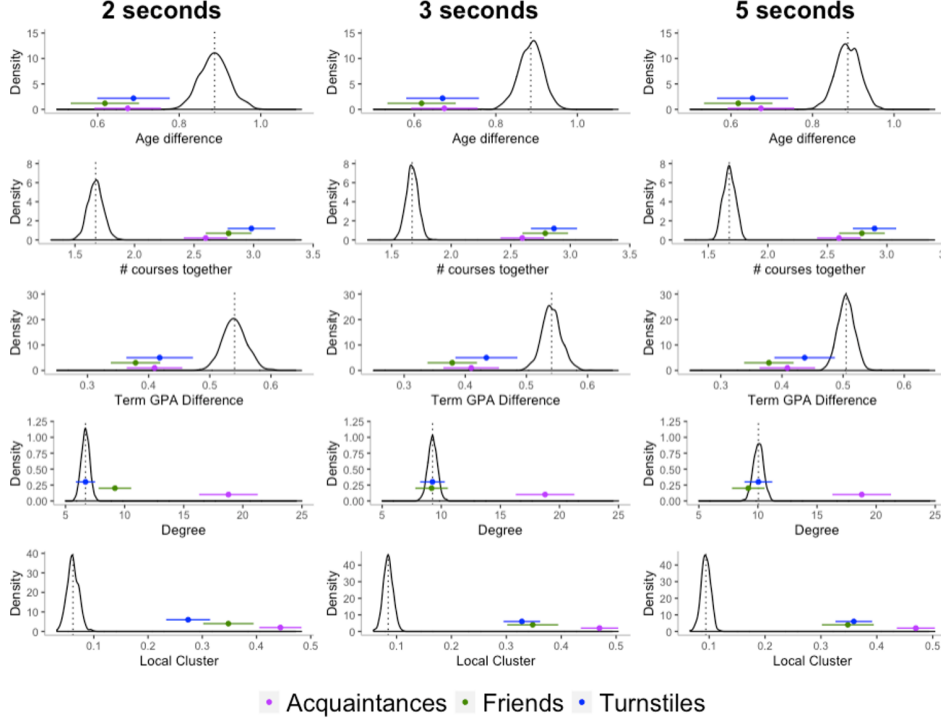
Measurement Error in Difference–in–Difference Framework: To build understanding about the role of measurement error in social interactions, I build on a potential outcomes framework in a 2x2 Difference–in–Difference research design as coined by Goodman–Bacon (2019) and discussed by Cunningham (2021). Define t as a treated group (i.e., a group with a large R_{mc}^l), and u as an untreated group:

$$\hat{\alpha_P}^{2x2} = (E[L_t|Post] - E[L_t|Pre]) - (E[L_u|Post] - E[L_u|Pre]) \tag{4}$$

In Equation 4, the estimated $\hat{\alpha_P}^{2x2}$ is written as the difference between the expected post– and pre–treatment value of the outcome L on the treated group t ($E[L_t|Post] - E[L_t|Pre]$), minus the difference between the expected post– and pre–treatment value of the outcome L on the untreated group u ($E[L_u|Post] - E[L_u|Pre]$). Equation 4 can be re-written in potential outcomes terms. Define L^0 as the potential outcome had no treatment be assigned and L^1 as the potential outcome had the treatment be assigned. Hence, the estimated $\hat{\alpha_P}^{2x2}$ can be re–written as:

$$\hat{\alpha_P}^{2x2} = \underbrace{E[L_t^1|Post] - E[L_t^0|Post]}_{\text{ATT}}$$

$$+ \underbrace{(E[L_t^0|Post] - E[L_t^0|Pre])}_{\text{Treatment counterfactual}} + \underbrace{(E[L_u^0|Post] - E[L_u^0|Pre])}_{\text{non-parallel trend bias==0}} \tag{5}$$

Equation 5 implies $\hat{\alpha_P}^{2x2}$ is made of the Average Treatment Effect on the Treated (ATT), which is the difference between the expected values of the outcome L on the post–treatment period and on the treated group t had the treated group received and not received the treatment, plus the non-parallel trend bias. The non–parallel trend bias is the difference in the potential outcomes for the treated and untreated group had no treatment be assigned to any group. I showed in Section 5 that there is no evidence of non–parallel trends bias. But, if the measurement error is associated with the treatment in ways unobserved by the researcher, then the estimated ATT based on the observed outcome may differ from the true ATT which I aim to estimate.

To fix ideas, define the number of links I aim to measure $L^{true} = L^{obs} - L^{F(+)} + L^{F(-)}$. That is, true links can be defined as the number of observed links L^{obs} minus the turnstile–elicited links which are false positives $L^{F(+)}$, plus the number of true links which were not captured by the turnstile–elicited measure $L^{F(-)}$ i.e., the false negatives. Then, the ATT I aim to estimate is:

$$ATT^{estimated} = E[L_t^{1,True}|Post] - E[L_t^{0,True}|Post] \tag{6}$$

Replacing $L_t^{1,True}$ and $L_t^{0,True}$ for their equivalent based on observed L, and doing some re-arraignment of terms I get:

$$ATT^{estimated} = E[L_t^{1,obs} - L_t^{1,F(+)} + L_t^{1,F(-)}|Post] - E[L_t^{0,obs} - L_t^{0,F(+)} + L_t^{0,F(-)}|Post]$$

$$= \underbrace{E[L_k^{1,obs}|Post] - E[L_k^{0,obs}|Post]}_{\text{Observed ATT}} +$$

$$\underbrace{E[L_t^{1,F(-)} - L_t^{1,F(+)}|Post] - E[L_t^{0,F(-)} - L_t^{0,F(+)}|Post]}_{\text{Measurement Error Bias}} \tag{7}$$

Thus, the estimated ATT can be re-written as the ATT based on the observed outcome L^{obs}, plus a measurement error bias, which can be described as the ATT on $L^{F(-)}$ minus ATT on

the $L^{F(+)}$:

$$ATT^{estimated} = ATT^{obs} + \underbrace{E[L_t^{1,F(-)} - L_t^{0,F(-)}|Post]}_{\text{ATT on F(-)}} - \underbrace{E[L_t^{1,F(+)} - L_t^{0,F(+)}|Post]}_{\text{ATT on F(+)}} \quad (8)$$

Equation 8 implies that if the treatment has no impact on $L^{F(-)}$ or $L^{F(+)}$ among the treated, then $ATT^{estimated} = ATT^{obs}$. In what follows, I discuss and test this implication in the context of my research design.

Ideally, I would have data on measurement error variables $L^{F(-)}$ and $L^{F(+)}$ across different majors and cohorts, in such a way that I can use variation in the treatment R_{mc}^l to assess its effects. Since I do not have data of that nature, I rely on proxy variables that can help me assess the extent to which the treatment R_{mc}^l may lead to measurement error in turnstile–elicited interactions. I use two variables to assess measurement error. First, the total number of ID taps at the turnstiles for each student. Second, I use the number of courses with turnstile–elicited links. I measure both proxies for the same terms I measure interactions (i.e., sixth and seventh terms after first enrollment).

Intuitively, if the treatment leads to more ID taps on the turnstiles the chances of capturing false positives $L^{F(+)}$ on the treated group increases. Similarly, if the treatment leads to fewer ID taps, the chances of missing true links $L^{F(-)}$ on the treated group increases. Likewise, treatment N_{Pmc} associated with a higher number of classes taken with the turnstile–elicited links may indicate higher chances of false positives $L^{F(+)}$. Classes in the sixth and seventh terms may be more diverse due to the treatment, but social interactions captured may be the product of chance. That is, wealthy students attending courses with other low–SES peers and coinciding in co–movements at the entrances, without that implying a true social interaction.

Table 9 displays the results of regressing R_{mc}^l on the measurement error proxies, under each time–window considered. The estimation follows the same structure as that of Equation 1 but using the proxy variables in the left–hand side. I do not find statistically significant evidence of a change in the number of ID taps at the turnstiles or in the number of courses with turnstile-elicited links due to exposure to low–income peers. Coupled with the previous results, I conclude there is no evidence to claim measurement error biases the estimated effects of exposure to low–income peers on students interactions or academic achievement.

Table 9: ATT on Measurement Error Proxies

		No. of courses with peers interacted		
	ID taps	2 seconds	3 seconds	5 seconds
	(1)	(2)	(3)	(4)
% of low–income peers	-4.162	0.003	-0.001	0.000
	(2.621)	(0.007)	(0.007)	(0.007)
Mu(Y)	1340	2.006	2.059	2.084
SD(Y)	1017	2.466	2.515	2.538

Notes: Results from estimating Equation **??** using measurement error proxies in the left–hand side. "ID taps" is the total number of ID taps of each students, either to enter or exit campus, in the sixth and seventh terms after first enrollment. "No. of courses with peers interacted" is the total number of courses the student took with the peers I defined as a turnstile–elicited link. All estimations include fixed effects by major and entry cohort as well as the covariates described for Equation **??**. All standard errors are clustered at the major-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Exposure to low–income in first–term courses

Exposure to low–income students operates not only through majors and cohorts but also in the classroom. Several of the courses offered to first–term students are open to multiple majors and offer multiple sections, which may lead to variations in exposure to low–income peers not accounted for by my initial difference–in–difference research design. For example, students at all the Engineering, Economics, and Business programs are required to take Differential Calculus in their first term. As a result, students in Business majors –which had virtually no change in the number of low–income students enrolled, may take Calculus with low–income peers from other majors, thus being exposed to low–income students in ways not accounted for by the entry cohort and major variation captured by Equation 1. Since this is the case for several other mandatory first term courses, the aggregated number of low–income peers a student may have in all their first–term courses may be much different than that captured by the major–cohort exposure utilized in the difference–in–difference strategy described so far.

To test for this source of bias, I exploit changes in the number of low–income peers wealthy students have throughout their first term courses and within majors and cohorts. In contrast to the difference–in–difference approach discussed this far, this design which does not rely on the quasi–random allocation of low–income students across majors and entry–cohorts, but on the change in low–income peers across all courses taken by a student and conditional on within major and entry cohort common shocks. To ensure the causal identification of the effects of individual–level exposure, I instrument the exposure to low–income peers across courses with its predicted number based on historical data on courses composition. I use the instrument to capture the exogenous variation in the size of exposure to low–income peers in first–term courses, had the distribution of low–income students not changed with the outset of SPP. This design allows me to assess the importance of individual level variations in exposure to low–income peers in explaining achievement and social interactions.

Equation 9 describes the specification I use to capture the effects of courses-based exposure to low–income peers. Y_i^{mc} represent the outcome of the relatively wealthy student i in a major and entry cohort group mc. The estimand of interest is ρ_l, which captures the response on the outcome to changes in the number of low–income peers IN_{imc}^l. Exposure to low–income peers may be driven by unobserved non–random aspects which are common to students in the same major and entry–cohort and which are related to academic achievement and may lead to differences in student friendships (number of courses required in the first term, size of courses, number of elective courses required, etc.). Thus, I include a major plus entry–cohort fixed effect ρ_{mc} which absorbs non-observed variation common to students

within each of these groups.

$$Y_i^{mc} = \rho_l IN_{imc}^l + \rho_N N_{imc}^s + \rho_s S_{imc} + \mathbf{X_i'}P + \rho_{mc} + \upsilon_{imc} \tag{9}$$

Importantly, SPP changed the number of low–income peers by increasing the total num-ber of students enrolled. To account for the change in the number of peers, I control for the total number of course peers across all courses the student took in their first term N_{imc}^s. This variable is computed like IN_{imc}^l, but without restricting the count by student SES back-ground. Exposure to peers is also relative to the number of classes each student is taking each term. Hence, I include as a control the number of courses taking in the first term of enrollment by student i (S_{imc}). Thus, random variation comes only through the number of low–income peers across courses and not through the change in class or cohort size. This approach is similar to that in Angrist and Lang (2004) and rationalizes the endogenous vari-ation through the number of low–income peers and conditional on the total number of peers. I condition my estimates on student characteristics represented by the $\mathbf{X_i}$ and discussed for Equation 1. Lastly, υ_{imc} represented the error term which is clustered at the major and entry–cohort level. Since overall courses selection in the first term may be driven by non-random unobserved preferences for peers, particularly in the SPP outset, I instrument N_{imc}^s with the predicted number of course–level peers.

I measure the exposure to low–income peers of each student IN_{imc}^P by counting the number of low–income students enrolled in each course each term, and by adding it for all courses taken by a student in their first term of college. For example, if a student is taking five courses in their first term and each course has four low–income students, then the value of the index for this students is twenty. This means I do not discount for repeated students. Rather, if the same low–income student shows up in multiple courses, I count them in as many courses as they show up as a way to incorporate the intensity of that exposure into my index. Figure 10 describes the change in the distribution of the index among relatively wealthy students in the entry cohorts before and after SPP. The values of the index change dramatically from the pre– to the post–SPP cohorts. The values of the index are all below 100 for the 2014 cohorts, and 75 percent of students do not have more than 34 low–income peers in their first term courses. In 2015, 25 percent of students have at least 148 low–income peers in their first term courses, with some students having over 300 low–income peers. Notwithstanding, the variation in the index should be considered in light of the identification strategy, which exploits the variation taking place within majors and entry cohorts.

Identification of the effect of exposure is challenged by students selecting in courses in ways associated with achievement or socializing preferences and not accounted by the research design. In particular, SPP students in the 2015 entry cohort were more likely to

Figure 10: Course-Level Exposure Index to low–income Peers in the First Term of Enrollment
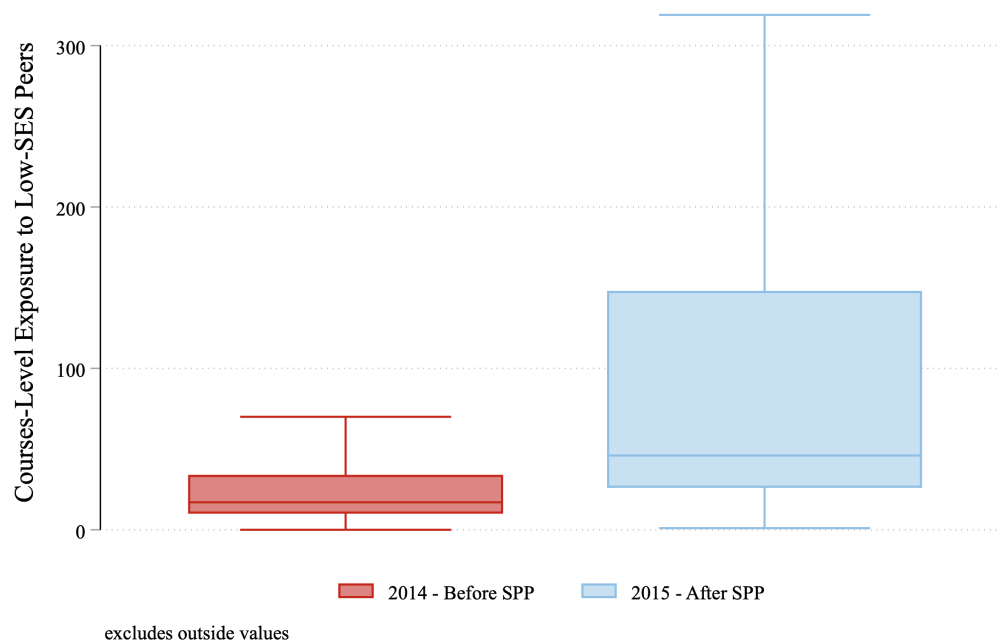
Notes: This figure plots the distribution of the course–level exposure to low–income peers in the first term index.
The plot follows the standard display of 75th percentile, median and 25th percentile references.

enroll in their courses late in the spring semester, as their overall enrollment process in the University was delayed due to the tight timing of the program, and relative to the traditional timing of wealthy students. As a result, low–income students may end up disproportionately allocated in certain courses of low–demand by other wealthy peers.

Since individual course–level exposure can be associated with unobserved non–random variation in the assignment of students to courses with the outset of SPP, I instrument it with the predicted allocation of low–income students to first term courses based on historical information from 2012 and 2013 course–level enrollment. Ultimately, I will test the effects of first–term course–level exposure to low–income peers on wealthy students achievement and social interactions using an instrumental variables approach that allows me to isolate exogenous variation in courses' socio–economic composition, in a Two Stage Least Square Fixed–Effects setting.

Computation of a Predicted Course–Level Exposure to low–income Peers: I use University data on course–level enrollment from 2012 and 2013 to estimate a distribution of low–income students across courses and to predict the number of low–income students in each course of the 2014 and 2015 cohort, had the distribution of low–income students not changed due to the outset of SPP. I computed the predicted distribution of low–income students as follows:

1. I keep the enrollment information for all courses offered between 2012 and 2013, one year before the start of SPP in 2015.

2. I keep the information only for students in their first–term of enrollment. I do this because I care about the allocation of incoming students across courses, and not about the allocation of other continuing students.

3. For each course –including all its sections and terms offered, I count the number of first–term low–income students enrolled.

4. I calculate the percentage distribution of first–term low–income students across all courses. I do this by dividing the number of low–income students in each course over the sum of low–income students across all courses. Doing so ensures that the sum of the percentage distribution of low–income students across courses adds to one.

5. Using the course–level data on first–term courses of the 2014 and 2015 entry cohorts, I match the distribution of low–income students per course computed in step 4. using the course ID information.

6. I predict the allocation of first–term students on those courses by multiplying the number of low–income students in each entry cohort with the estimated percentage distribution of the course.

7. For each student, I add the predicted allocation of low–income peers in their first term courses. This gives the predicted index of exposure to low–income peers in the first term courses, had the allocation of low–income students not changed with SPP.

The instrument show a high relevance as stated by the first-stage estimates displayed in Tables 10 and 11 which display the estimated effects on number of links and academic achievement. Specifically, the point estimate of the first stage is statistically significant with a 99 percent confidence, indicate a positive relation and a sizable point estimate. Following Angrist and Pischke (2009), these tables also include the F–test of Excluded Instruments which is above 20 across all specifications, which rejects the null hypothesis of a weak instrument.

Table 10: Effect of Courses–level Exposure to low–income Peers on Wealthy Students' Friendships

	2 - seconds window			3 - seconds window			5 - seconds window		
	(1) Any	(2) Wealthy	(3) Low SES	(4) Any	(5) Wealthy	(6) Low SES	(7) Any	(8) Wealthy	(9) Low SES
2SLS									
IN_{imc}	0.001	-0.002	0.003	0.007	-0.001	0.008	0.014	0.003	0.011
	(0.018)	(0.015)	(0.006)	(0.024)	(0.019)	(0.008)	(0.025)	(0.020)	(0.008)
First Stage									
Predicted IN_{imc}	0.460***			0.460***			0.460***		
	(0.102)			(0.102)			(0.102)		
F-test excluded instruments	20.28	20.28	20.28	20.28	20.28	20.28	20.28	20.28	20.28
Reduced Form									
Predicted IN_{imc}	0.001	-0.001	0.002	0.003	-0.000	0.004	0.006	0.001	0.005
	(0.009)	(0.007)	(0.003)	(0.011)	(0.009)	(0.004)	(0.012)	(0.009)	(0.004)
OLS									
IN_{imc}	-0.000	-0.002	0.002	-0.000	-0.003	0.003**	0.001	-0.003	0.003*
	(0.003)	(0.002)	(0.001)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)	(0.002)
Pre-treatment statistics									
mu(No. of friends)	4.085	3.892	0.193	5.212	4.973	0.239	5.002	4.771	0.231
sd(No. of friends)	4.213	4.019	0.494	5.154	4.922	0.558	5.040	4.816	0.548
sd(IN_{imc})	34.31	34.31	34.31	34.31	34.31	34.31	34.31	34.31	34.31
sd(predicted)	7.231	7.231	7.231	7.231	7.231	7.231	7.231	7.231	7.231
No. Students	5,278	5,278	5,278	5,278	5,278	5,278	5,278	5,278	5,278

Notes: This table displays the estimates from Equation 9. All estimates include fixed effects by major–cohort. IN_{imc} is a measure of individual–level exposure to low–income peers in the first term of college based on the courses students take in that term. The computation of IN_{imc} and the predicted IN_{imc} used as instrument is described in Appendix B. All standard errors are clustered at the major-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Effect of Courses-Level Exposure to low–income Peers on Wealthy Students' Achievement

	(1)	(2)	(3)	(4)	(5)	(6)
	6th term cum. Credits	6th term cum. GPA	7th term cum. Credits	7th term cum. GPA	Graduated in 8 term	Graduated in 9 terms
2SLS						
IN_{imc}	0.023	-0.001	-0.003	-0.001	-0.002	0.001
	(0.071)	(0.001)	(0.079)	(0.001)	(0.002)	(0.002)
First Stage						
Predicted IN_{imc}	0.511***	0.511***	0.533***	0.533***	0.460***	0.574***
	(0.110)	(0.110)	(0.112)	(0.112)	(0.102)	(0.102)
F-test excluded instruments	21.74	21.74	22.52	22.52	20.28	31.92
Reduced Form						
Predicted IN_{imc}	0.012	-0.001	-0.002	-0.001	-0.001	0.000
	(0.036)	(0.001)	(0.043)	(0.001)	(0.001)	(0.001)
OLS						
IN_{imc}	-0.033***	-0.001**	-0.043***	-0.001***	-0.000	-0.000
	(0.012)	(0.000)	(0.013)	(0.000)	(0.000)	(0.000)
Pre-treatment statistics						
mu(Y)	100.9	3.859	116.6	3.870	0.0693	0.228
sd(Y)	16.22	0.344	18.95	0.340	0.254	0.420
sd(IN_{imc})	34.31	34.31	34.31	34.31	34.31	34.31
sd(predicted IN_{imc})	7.231	7.231	7.231	7.231	7.231	7.231
No. Students	4,507	4,507	4,447	4,447	5,278	4,027

Notes: This table displays the estimates from Equation 9. All estimates include fixed effects by major–cohort. IN_{imc} is a measure of individual–level exposure to low–income peers in the first term of college based on the courses students take in that term. The computation of IN_{imc} and the predicted IN_{imc} used as instrument is described in Appendix B. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.

C   Appendix: Effects of Exposure on Social Interactions Under Alternative Definitions of Turnstile–elicited Links

Table 12: The Impact of Exposure to Desegregation on the Links of Wealthy Students - 2 seconds windows

	A. Number of Links			B. Probability of a Link with		C. Proportion of Links
	(1)	(2)	(3)	(3)	(4)	(5)
L_imc with:	Any	Wealthy	Low Income	Wealthy	Low Income	Low Income
OLS						
% of low–income peers	-0.013	-0.036***	0.023***	-0.003**	0.007***	0.707***
	(0.013)	(0.013)	(0.003)	(0.001)	(0.001)	(0.070)
B. Non-linear Effects						
$\mathbb{I}$[% of low–income peers > 30%]	-0.097	-0.568	0.470***	-0.042	0.111***	15.659***
	(0.412)	(0.440)	(0.100)	(0.037)	(0.032)	(3.045)
Pre-treatment statistics of the Outcomes						
mean	4.085	3.892	0.193	0.736	0.157	4.330
standard deviation	4.213	4.019	0.494	0.441	0.364	11.51
Treatment distribution in 2015						
$\Delta \mathrm{mu}(R^l_{mc})$			9.51 p. p.			
$\mathrm{sd}(R^l_{mc})$			12.98			
No. Students	5,278	5,278	5,278	5,278	5,278	3,979

Notes: This table displays the estimates from Equation 1 in Panel A., and the non–linear estimates in Panel B. Outcomes are based on a turnstile–elicited links based on time–windows of 2 seconds and at least two co–movements in a term. *% of low–income peers* is calculated at the major–cohort levels. $\mathbb{I}$[% of low–income peers > 30%] is an indicator function equal to one if the share of low–income peers in the major and cohort is greater than 30% i.e., over the 75th percentile of the distribution in the 2015 – Spring entry cohort. All estimations control for female indicators, age in year at the time of entry, SB11 standardized test score, mothers' has no college education indicator, indicator of middle–SES background according to the social strata indicator, and dummies for whether the student got an SPP loan. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: The Impact of Exposure to Desegregation on the Links of Wealthy Students - 5 Seconds windows

	A. Number of Links			B. Probability of a Link with		C. Proportion of Links
	(1)	(2)	(3)	(3)	(4)	(5)
L_imc with:	Any	Wealthy	Low Income	Wealthy	Low Income	Low Income
OLS						
% of low–income peers	-0.010	-0.043***	0.033***	-0.002**	0.008***	0.706***
	(0.017)	(0.016)	(0.004)	(0.001)	(0.001)	(0.050)
B. Non-linear Effects						
$\mathbb{I}$[% of low–income peers > 30%]	-0.001	-0.775	0.774***	-0.045	0.118***	15.699***
	(0.515)	(0.497)	(0.144)	(0.029)	(0.038)	(2.610)
Pre-treatment statistics						
mu(Outcome)	5.002	4.771	0.231	0.752	0.180	4.386
sd(Outcome)	5.040	4.816	0.548	0.432	0.384	11.14
Treatment distribution in 2015						
Δ mu(R^l_{mc})			9.51 p. p.			
sd(R^l_{mc})			12.98			
No. Students	5,278	5,278	5,278	5,278	5,278	4,063

Notes: This table displays the estimates from Equation 1 in Panel A., and the non–linear estimates in Panel B. Outcomes are based on a turnstile–elicited links based on time–windows of 5 seconds and at least two co–movements in a term. *% of low–income peers* is calculated at the major–cohort levels. $\mathbb{I}$[% of low–income peers > 30%] is an indicator function equal to one if the share of low–income peers in the major and cohort is greater than 30% i.e., over the 75th percentile of the distribution in the 2015 − Spring entry cohort. All estimations control for female indicators, age in year at the time of entry, SB11 standardized test score, mothers' has no college education indicator, indicator of middle–SES background according to the social strata indicator, and dummies for whether the student got an SPP loan. I address missing values by including an indicator equal to 1 if any of the covariates contains a missing value for the student i. All standard errors are clustered at the major-cohort level. * p < 0.10, ** p < 0.05, *** p < 0.01.