“Classroom composition and network effects: Evidence from a college special admission program in Chile”

Nicolas Suarez Chavarria

September 19, 2021

Research idea

Using administrative data about the grades of Business and Economics undergraduate students from Chile, and exploiting the fact that students are randomly assigned to their first semester classes, I want to examine the existence of peer effects among students, and to study how different class compositions affect their outcomes. In 2012, the University of Chile implemented a college special admission program that targets high achieving vulnerable students from public schools, and I want to see how students that belong or don't belong to this program interact within their own group and with the other group, and how students should be assigned to classes in order to maximize their academic outcomes.

Besides the fact that in this context the peers of the students are determined in an exogenous way, another advantage of this setup is these college students are assigned on average to 6 classes during their first semester, and they interact with different students in each of these classes, so this allows me to use instrumental variables in a similar fashion to Bramoulle et al. (2009), computing measures of the outcomes or characteristics of the peers of the peers of a student, in contrast to other type of studies where all students belong to the same peer group, as it is the case with roommate or classroom studies. Besides that, the network structure can be exploited to compute network characteristics that could be able to explain how the interactions between these groups might determine the outcomes of the students in this context.

Motivation

S. Zimmerman (2019) studied how elite colleges help students reach top positions in the economy in Chile. He identified 3 highly selective, business-oriented majors dictated at the 2 most selective universities in Chile, and found that 1.8% of their graduates account for 41% of leadership positions in major companies and the same share accounts for 39% of top 0.1% incomes. One of these programs is the Business Engineering program at the University of Chile, for which I have administrative data, so I want to study what happens when vulnerable students from public high schools interact in the classroom with high achieving students from expensive elite high schools, and how these students affect people in their same group, and how exposure to people from a different group can affect them.

In this context, besides estimating peer effects on their own, another objective is to study how students *should be* assigned to classes in order to maximize their outcomes. We can look at classroom compositions and the overall structure of our peer's network to see what are desirable traits in a design either for the students that are part of the affirmative action program, or for all the members of the student body. The affirmative action implemented in this school was also implemented in other schools at the University of Chile, so what we learn here could have educational policy implications for an important number of students.

Context

In Chile there is a centralized admission system through which students apply to multiple programs and universities. In order to apply to the vast majority of accredited programs and universities, students need to take the "University Selection Test" ("Prueba de Seleccion Universitaria" in Spanish or **PSU**)[1], a battery of 4 standardized tests to test their competencies in Mathematics, Spanish, Natural Sciences and History. After knowing their scores, students apply to several programs through a unique platform, submitting a ranked list of preferences over up to 10 pairs of program/universities. Students and their applications are ranked by universities based on a weighted average of their PSU scores, high school GPA and high school within-cohort ranking, with weights that are specific to each program. After ranking the applicants, universities select students based on their weighted scores using a deterred acceptance algorithm.

In this context, during the year 2012 the University of Chile implemented the **SIPEE** program (Priority Entry System for Educational Equity, or "Sistema de Ingreso Prioritario de Equidad Educativa" in Spanish), a special admission mechanism to admit high achieving students from public high schools that applied to programs at the university but were below the admission cutoff score. These students have to comply with the following requisites to be able to apply through this special mechanism:

- Taking their university selection test the same year they graduate from high school.
- Having completed all their high school education in a public school.
- Studying during their senior high school year in a high school with a vulnerability score (how many of their students are socioeconomically vulnerable) above 30%.
- Being below the 60th percentile in the Social Household Registry (the registry ranks household mostly according on their per capita income).
- Having a GPA above 5.5 (in a scale from 1 to 7) during all high school.

The SIPEE program started offering 131 reserved seats, but has been scaled up since 2012, and now offers 500 reserved seats per year, with 50 of them belonging for programs in the School of Economics and Business. The seats are filled by the eligible students below the admission cutoff score, ranking them on their score.

At the moment, I'm evaluating this program together with the supernumerary seat program for **BEA**[2] students. This program guarantees seats to students vulnerable students from public schools within the top 10% of their promotions that received a scholarship from the government, but that are below the admission cutoff score. These seats are also assigned to eligible students ranked based on their admission scores. During 2012 the School of Economics and Business offered 30 supernumerary seats for BEA students, but they increased the number of seats during

[1]Some artistic programs require special tests that replace the PSU

[2]BEA stands for "Beca de excelencia academica" in Spanish, which can be translated as "Scholarship for Academic Excellence"

time, reaching 45 during 2020.

I'm evaluating together the SIPEE and BEA programs since there is a significant overlap between their eligibility criteria, so in the end both programs target similar groups of students.

I'm going to study the peer effects within and between groups for students that were admitted through regular admission and students of the SIPEE and BEA programs using data for the 3 careers imparted at the School of Economics and Business:

- **Business engineering**: 5 year program that leads to a bachelor degree in Economics or Business. Accounts for approximately 70% of enrollment in the School of Economics and Business, and it is the most selective of the 3 careers (it has the highest admission cutoff score every year).
- **Information and Control Management Systems engineering**: 5 year program that accounts for approximately 25% of enrollment.
- **Accounting**: 5 year program that accounts for approximately 5% of enrollment.

During the 2012-2020 period, the 3 careers in the School of Economics and Business have undergone 2 reforms in their academic curriculum, with the first curriculum being valid until 2012, the second one was current between 2013 and 2019, and the third being implemented during 2020. There are some common characteristics in these 3 curricula: during the first semester, students always take between 5 and 7 classes depending on their curriculum, one of them being an English class. For their English classes, students take an ETS' TOEIC English test at the moment of enrolling to determine their English proficiency, and based on that they are assigned to a certain class. Students of all 3 careers are then randomly assigned to English classes conditional on their English proficiency.

For the rest of the classes, which generally include Algebra, Business and Economics introductory classes, students from the 3 careers are randomly assigned to classes. Almost all the classes are common across all 3 careers for all curricula, so students can be assigned to classes where students from other careers may be enrolled, but there are 1 or 2 classes that are career-specific in some of the curricula.

Related literature

There is a wide body of literature concerned with peer effects in general, starting with Manski (1993) and followed later by Bramoulle et al. (2009) and others. When looking at peer effects in an educational context, we have articles that study peer effects in school classrooms without random assignment to classes, like Hoxby (2000), McEwan (2003), Hanushek et al. (2003) and Hoxby and Weingarth (2005). There are also several studies focusing on peer effects associated to roommates in colleges , like Sacerdote (2001), D. Zimmerman (2003) and Kremer and Levy (2008).

There are also several papers that study educational peer effects in contexts with random assignment: Whitmore (2005) studies the effect of gender composition of classroom over the achievement of students using Tennessee's project STAR data, and finds that being in a classroom where the majority of the students are women has a positive impact over achievement, but the size of the effect decreases over time, and by the time students are in grade 3, the effect turns negative. Ammermueller and Pischke (2006) estimates peer effects for fourth graders in 6 European countries, finding positive peer effects overall, but there is significant heterogeneity across countries. Kang (2007) finds moderate yet positive peer effects in the math scores of middle school students in South Korea, using an instrumental variables approach. We also have Carrell et al. (2009), a study where they look at peer effect for college students that are randomly assigned to groups that live nearby and interact among themselves heavily.

There are not a lot of studies with college data and randomly assigned peers, and to the extend of my knowledge there are no articles with a peer structure quite the one presented in this paper.

Empirical strategy

Given the nature of my data, a very simple starting point is to estimate the effect of classroom composition and network structure over the outcomes of the student at the class level. To do this, I estimate the following equation:

$$y_{i,c} = a_i + CK_{i,c}\beta_1 + CK_{i,c} \times S_i + \varepsilon_{i,c} \tag{1}$$

Where $y_{i,c}$ is the outcome of student i in class c, a_i is a student level fixed effect, CK_c is a classroom composition variable (measured either as the share SIPEE students in the classroom or as the homophily for student i in the classroom[3]), and S_i is an dummy indicating if student i is part of the SIPEE program or not. The idea with this equation is that our student level fixed effect should be a very strong predictor of the outcomes of a student, so if we find an effect of class composition over outcomes even after this, that would provide initial evidence that the interactions between individuals in our 2 groups are important.

Next I study the effect of some characteristics of the network over the outcomes at the student level:

$$y_i = a_i + f(N, S_i)\beta + \varepsilon_{i,c} \tag{2}$$

Where y_i is an outcome of student i over several classes, $f(N, S_i)$ is some network characteristic for student i. Here I use the degree of student i (the number of peers she has) and their clustering coefficient (the share of her peers that are also peers among themselves). I also compute these 2 measures for students within each of the 2 groups, so I can obtain differential effects between groups. This equation can help us design classrooms: we can see how beneficial for a student is to belong to a group that interacts in several contexts, if there are benefits to have very closed peer circles within their own group, and to see how clustered are my peers in the other group

[3] Defined as the share of students in the same group (SIPEE or non-SIPEE) as student i.

could affect me. It would be plausible to think that having a very clustered peer circle of people in your own group would be a good thing, since students would spend a lot of their time together, maybe form study groups or share notes. Regarding the peers on the other group, it is probably more beneficial for a student if they are not very clustered, since the other group being very close would make it harder for the student to interact with them.

After that, I want to look for the existence of peer effects in our sample. To do this, we can follow Bramoulle et al. (2009) and De Giorgi et al. (2010), and estimate the following equation (in matrix form):

$$y = \alpha + Gy\beta + X\gamma + GX\delta + \varepsilon \tag{3}$$

Where y is a $N \times 1$ vector of outcomes, X is a $N \times K$ matrix of covariates, G is a $N \times N$ adjacency matrix and ε is an $N \times 1$ error vector. For our adjacency matrix G we can define element G_{ij} as the number of times individuals i and j took a class together over the total number of interactions that individual i had. We are also going to define $G_{ii} = 0$ for all $i \in N$. Given this, G is a block diagonal right stochastic matrix. In this context, β is capturing endogenous peer effects and δ captures exogenous peer effects.

Bramoulle et al. (2009) showed that this model is identified as long as I, G and G^2 are linearly independent. De Giorgi et al. (2010) does something similar, studying the effect of peers over major choice in college in Italy, also defining peers based on randomly assigned first semester classes, and relying on the existence of excluded peers (students that are peers of the peers of student i, but that are not peers of i themselves) and a 2SLS strategy to identify peer effects. Here, we are going to do something similar to estimate equation (3). In this context, we also need our matrix G to be exogenous: conditional on their English level, students in my sample are randomly assigned to first semester classes, so their peers are indeed exogenous.

Data

I use administrative data for the 2012 to 2020 cohorts of the 3 programs imparted at the School of Economics and Business at the University of Chile. The data includes information about gender, birth date, program, academic situation, admission system, admission score, commune of residency, current GPA[4] and classes taken during their first semester. I'm currently dropping from my sample transferred students and students from other special admission programs (sport scholarships, international students and others), and pooling together students from the SIPEE and BEA programs. In Table 1 I summarize the number of students by cohort and career: we can see that Business Engineering accounts for most of the enrollment per cohort, and that enrollment has increased over time for all 3 careers.

[4]This GPA includes all passed and failed classes during the complete career of graduated students, and the current GPA for students that were still enrolled during 2020

Table 1: Students by career

Year	**Program**		
	Business Engineering	**Information Engineering**	**Accounting**
2012	382	128	19
2013	379	102	40
2014	396	122	66
2015	397	143	69
2016	403	142	73
2017	417	158	83
2018	452	172	80
2019	481	181	95
2020	469	184	97
Total	3,776	1,332	622

In Table 2 I separate enrollment by type of admission, where we can that the SIPEE and BEA students represent in every cohort roughly 10% of enrollment.

Table 2: Students by type of admission

Year	**Type of admission**	
	SIPEE & BEA	**Regular**
2012	51	478
2013	69	452
2014	79	505
2015	66	543
2016	79	539
2017	62	596
2018	78	626
2019	70	687
2020	59	691
Total	613	5,117

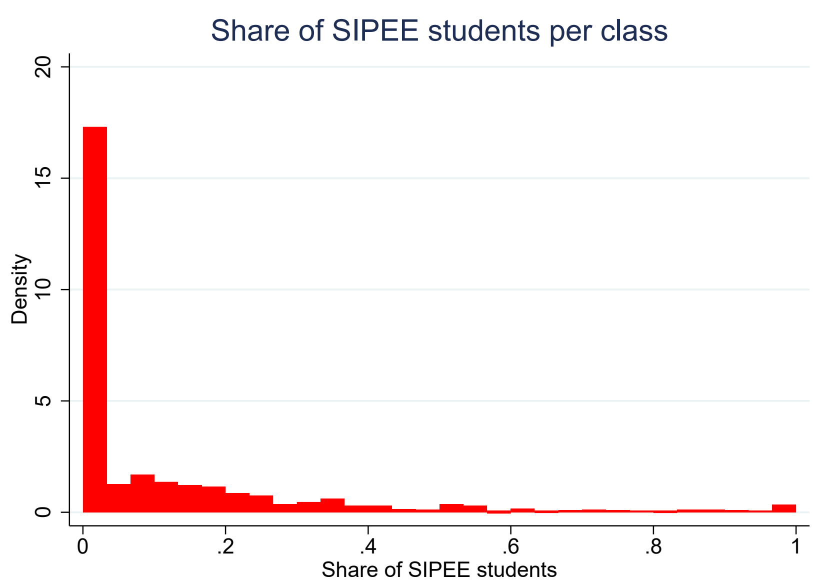
Next, in Table 3 we can see some descriptive statistics for our sample, observing values for a student, all their peers, their peers in the SIPEE program and outside the SIPEE program: 40% of students are women, and on average 40% of the peers of a student are women, without important differences between SIPEE and non SIPEE peers. As we saw before 11% of students are part of the SIPEE program, and the all students face the same share of SIPEE peers. Regarding admission scores, by construction we have that SIPEE peers have a lower admission score, since all the students in the program were admitted with scores below the admission score cutoff point. Our GPA measure is standardized, so it is no surprise that the average GPA is 0 with a standard deviation of 1. The peers of a student have a similar GPA but with less dispersion, and SIPEE peers have a slightly higher GPA than their non-SIPEE counterparts. When we look at the average grades of the first semester only, we see similar patterns, since these grades were also standardized, but now we see that non-SIPEE peers perform better than SIPEE peers. Finally, the passing rate of first semester classes here is around 90%, and we see that SIPEE peers have significantly lower passing rates than the rest of the peers.

Table 3: Student level descriptive statistics

Variable	**Own student**		**Weighted mean of peers**					
			All		**SIPEE**		**non-SIPEE**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Female	0.40	0.49	0.40	0.04	0.43	0.19	0.39	0.04
SIPEE student	0.11	0.31	0.10	0.11	-	-	-	-
Admission score	701.64	31.20	701.82	10.04	596.02	204.53	664.17	116.45
Overall GPA	0.01	0.98	0.02	0.09	0.02	0.24	0.01	0.12
First semester GPA	0.02	0.68	0.02	0.07	-0.11	0.15	0.03	0.08
Passing rate	0.89	0.17	0.89	0.03	0.76	0.23	0.90	0.03

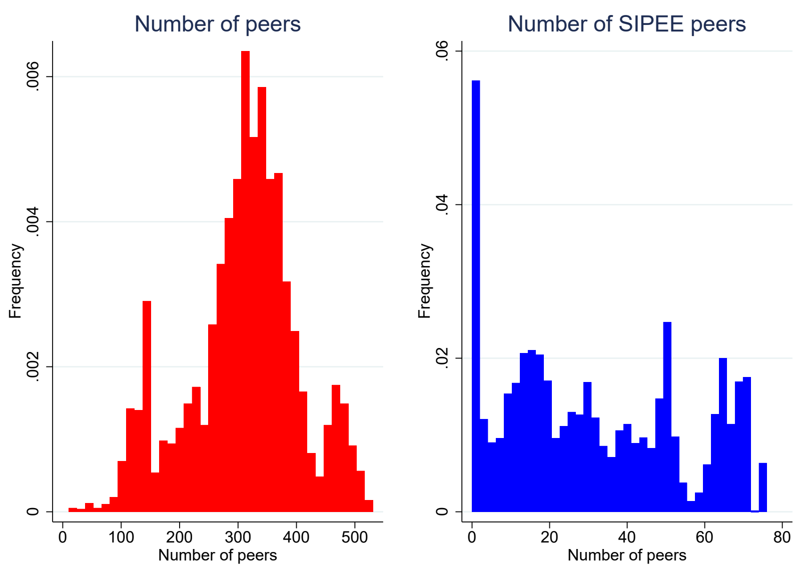
Now we can focus on describing the peer network: in Figure 1 we can see the distribution of the share of SIPEE students within each of the 1,229 classes present in our dataset. There is a lot of variation in SIPEE shares: while around 55% of all the first semester are completely made of non-SIPEE students, in the remaining 45% we observe almost all possible compositions. In fact, there were even 13 classes were all the students were part of the SIPEE program.

Figure 1: Distribution of SIPEE share per class



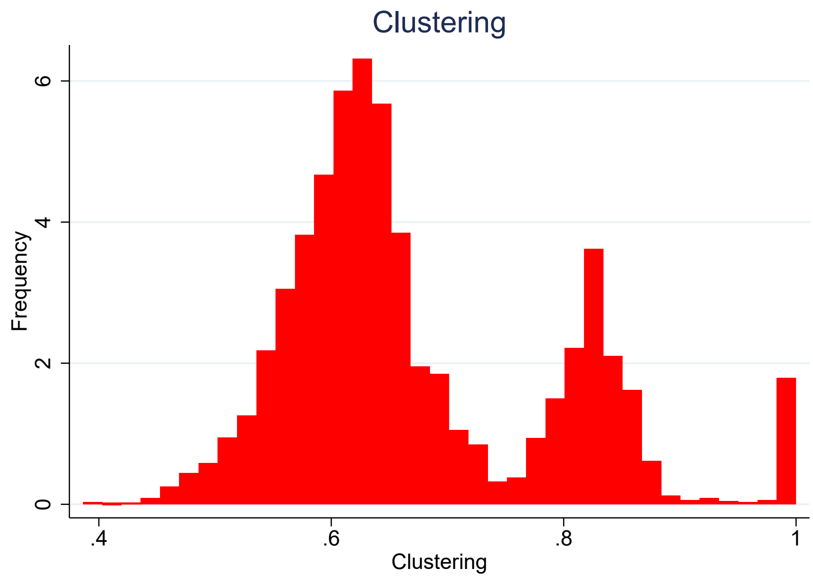
Next in Figure 2 we can see the distribution of the degree of the students, where I counted the links rather than adding the intensity of the links. We observe that most of the students have a substantial number of peers, with most of the mass of students having around 300 peers, and no students without peers (which is to be expected, since all students take classes with at least 1 other person in them). If we look at the number of SIPEE peers, now we see that a considerable number of students don't have SIPEE peers, but some of them are exposed to almost 80 SIPEE peers. This was to be expected, since the SIPEE students per cohort are always around 10% of the total student body.

Figure 2: Distribution of the degree of students



In Figure 3 we can now observe the individual clustering coefficients of all the students in our sample. Here I also ignore the intensity of the link between students, and only considered if they are or not connected. Our clustering coefficient indicates what share of my peers are also peers themselves. We observe a wide range of clustering among individuals, with most students having around 60% of their peers being peers themselves.

Figure 3: Distribution of student's clustering



Finally, in Table 4 we see more detailed information about the network characteristics by group. The average student has around 310 peers and has a clustering coefficient of 67%. There is not much of a point in analyzing the complete sample, since the statistics there are mainly driven by the non-SIPEE students. Non-SIPEE students have on average more peers that their SIPEE counterparts, they obviously have more peers in their same group than in the other group since they are the dominant group in terms of size. It is interesting to note that even if on average SIPEE students have less peers, they also generally have more peers in the SIPEE group compared to the non-SIPEE students (56 versus 28 on average). Regarding clustering, we can see that SIPEE students are a little more clustered than non-SIPEE students, and surprisingly, for both groups it is true that the clustering of their peers in the other group is bigger than the clustering of their peers in their same group. It is very interesting to see that the SIPEE peers of SIPEE students have on average a very low clustering coefficient.

Table 4: Network characteristics by group

Variable	Mean	Std. Dev.	Minimum	Median	Maximum
All observations:					
Number of peers	310.22	94.24	10.00	318.00	531.00
Peers in the same group	259.22	107.93	10.00	288.00	498.00
Peers in the other group	51.00	72.35	0.00	29.00	393.00
Clustering	0.67	0.12	0.39	0.64	1.00
Clustering with the same group	0.61	0.16	0.04	0.62	1.00
Clustering with the other group	0.79	0.30	0.00	0.93	1.00
Non-SIPEE students:					
Number of peers	313.01	94.61	10.00	324.00	531.00
Peers in the same group	284.35	85.81	10.00	299.50	498.00
Peers in the other group	28.67	21.96	0.00	24.00	76.00
Clustering	0.67	0.11	0.39	0.64	1.00
Clustering with the same group	0.66	0.11	0.39	0.63	1.00
Clustering with the other group	0.80	0.31	0.00	0.96	1.00
SIPEE students:					
Number of peers	287.74	88.18	44.00	283.00	463.00
Peers in the same group	56.93	11.87	11.00	61.00	75.00
Peers in the other group	230.81	84.58	3.00	228.00	393.00
Clustering	0.70	0.14	0.48	0.64	1.00
Clustering with the same group	0.27	0.12	0.04	0.25	0.92
Clustering with the other group	0.68	0.15	0.48	0.64	1.00

Preliminary results

In Table 5 I report the estimates of equation (1). I study the effect of class composition over class level outcomes: the outcomes here are the grades obtained by students in each class, and whether they passed a class or not. All regressions include student level fixed effects. I use 2 measures of class composition: the share of SIPEE students in a class, and homophily, defined as the share of students that are in the same group as the students (so it is the SIPEE share for SIPEE students, and the non-SIPEE share for the rest of the students). For each composition variable, I estimate an equation with only that variable, and then I add an interaction with the SIPEE dummy. In panel A of Table 5 we observe the SIPEE share results: we see a positive and statistically significant effect of the SIPEE share over grades, and no interaction effect. Regarding passing rates, we see no effect of the SIPEE share on its own, but we see a net positive effect of SIPEE share over SIPEE students, and a negative effect of the SIPEE share for non-SIPEE students. So here increasing the share of SIPEE students in a classroom would improve grades for everyone, but at the same time it would make SIPEE students more likely to pass, and non-SIPEE students more likely to fail.

Now we look at the effects of homophily in the classroom in panel B: there is not an statistically significant effect over grades in the model without interactions, but once we add the interaction we see again a net positive effect for SIPEE students (so they benefit from having more SIPEE students in the classroom), and a negative effect for non-SIPEE students, so they are worse off when more non-SIPEE students are they peers. This is consistent with the positive effects in the SIPEE share equations of panel A. If we look at the passing rates, we now see a positive effect of homophily over passing rates, and no differential effects. This is again consistent with the results in panel A: for SIPEE students, having more SIPEE students is better, while for non-SIPEE students, it would be beneficial to have less SIPEE students in the classroom and more of their own.

Table 5: Effects of class composition over per class outcomes

Variable	**Grade**				**Passing**			
Panel A: SIPEE share								
Share of SIPEE students in the class	0.2582	[0.00]	0.2522	[0.00]	-0.0229	[0.36]	-0.0914	[0.00]
	(0.041)		(0.054)		(0.025)		(0.029)	
SIPEE × Share of SIPEE students			0.0175	[0.86]			0.2001	[0.00]
			(0.097)				(0.060)	
Panel B: Homophily								
Homophily	-0.0741	[0.12]	-0.2522	[0.00]	0.0973	[0.00]	0.0914	[0.00]
	(0.048)		(0.054)		(0.027)		(0.029)	
SIPEE × Homophily			0.5220	[0.00]			0.0173	[0.75]
			(0.087)				(0.055)	
Observations	31,393		31,393		31,434		31,434	

Note: Model includes student level fixed effect. P-values are reported in squared brackets.
Double clustered at the student and class levels standard errors are reported in parentheses.

Even if we found statistically significant effects even after having student level fixed effects, and we have results that are consistent between the 2 panels, there is still no clear evidence on what kind of composition would be the best in the classroom, specially because the results imply that non-SIPEE students improve their grades, yet they are more likely to fail their classes, which make no logical sense.

Next, in Table 6 we estimate equation (2) and study student level effects of network characteristics over outcomes. Here the possible outcomes are the GPA of the student during their first semester, the overall GPA of the student during their whole stay in the university and the passing rate of the student for their first semester classes. This model includes controls for the admission score of the student, the weighted average admission score of the peers, and cohort level fixed effects. As network characteristics we use the degree of the student (number of peers) and their clustering coefficient. In on specification we use these overall network characteristics, but in the other we separate them per group.

Table 6: Effects of network characteristics over student outcomes

Variable	1st semester GPA				Career GPA				Passing rate			
Panel A: Degree												
Number of peers	-0.0001	[0.32]			-0.0003	[0.17]			-0.0000	[0.92]		
	(0.000)				(0.000)				(0.000)			
Number of peers in own group			-0.0001	[0.46]			-0.0001	[0.58]			-0.0000	[0.67]
			(0.000)				(0.000)				(0.000)	
Number of peers in other group			0.0000	[0.92]			0.0011	[0.01]			-0.0001	[0.09]
			(0.000)				(0.000)				(0.000)	
Panel B: Clustering												
Clustering among peers	-0.0200	[0.83]			-0.1215	[0.22]			0.0071	[0.84]		
	(0.094)				(0.099)				(0.034)			
Clustering in own group			-0.0906	[0.27]			-0.6129	[0.00]			0.0482	[0.02]
			(0.082)				(0.091)				(0.021)	
Clustering in other group			-0.0117	[0.68]			-0.0027	[0.95]			-0.0059	[0.13]
			(0.028)				(0.046)				(0.004)	
Observations	5,128		5,128		5,126		5,126		5,128		5,128	

Note: Model includes controls for the admission score of the student, the average admission score of the peers, and cohort level fixed effects.
P-values are reported in squared brackets. Standard erros clustered at the cohort level are reported in parentheses.

In panel A of Table 6 we observe the effects of the degree of a student. I find no effect of the number of peers over the first semester GPA with the 2 specifications, there is no effect of the overall number of peers over the career GPA, but there is a very small positive effect of the number of peers in the other group over the career GPA: to increase their career GPA in one standard deviation, a student would need to have more than 900 extra peers in the other group, which is impossible given the cohort sizes. Finally, there is only an extremely small negative effect of the number of peers in the other group over passing rates, but this is not relevant given the magnitude of the effect.

On the other hand, in panel B of Table 6 we find very different results: There are again no effects over the first semester GPA, there is a negative effect of the clustering with peers in the same group of the student over the career GPA, but there is a positive effect of the same variable (although much smaller in magnitude) over the passing rates of students. The evidence in these 2 panels is not necessarily inconsistent, since it is possible for a variable to have opposite effects over GPA and passing rates, specially since the career GPA and the passing rate of the first semester are defined over very different time horizons. But since the signs are opposites, there is not an objective answer to determine with what kind of peers should a student in a group interact, or what degree of clustering these groups should have.

In Table 7 we observe the estimates of equation (3) using OLS and student level data. Here again the possible outcomes are the first semester GPA of a student, the GPA of the student over all his time studying (some cohorts have graduated, some are still studying), and the passing rate of the classes taken during the first semester. I control for gender, admission type and admission scores. In the even columns I add cohort level fixed effects.

If we look at effects over the first semester GPA we find a positive endogenous effect only when we include cohort level fixed effects. In both specifications there is a positive and statistically significant effect of being a woman over GPA, there are no differences caused by admission type, and admission scores have a positive effect. Regarding exogenous effects, we only find a negative and significant effect of admission scores over GPA when we add cohort fixed effects.

Table 7: Peer effects OLS estimates

Variable	**1st semester GPA**				**Career GPA**				**Passing rate**			
Outcome of peers	0.0710	[0.71]	0.3711	[0.02]	-0.0318	[0.83]	0.2970	[0.01]	0.9084	[0.00]	-0.0583	[0.81]
	(0.189)		(0.162)		(0.151)		(0.118)		(0.072)		(0.239)	
Female	0.2082	[0.00]	0.2047	[0.00]	0.2083	[0.00]	0.2054	[0.00]	0.0191	[0.00]	0.0185	[0.00]
	(0.018)		(0.018)		(0.018)		(0.018)		(0.005)		(0.005)	
SIPEE student	0.0632	[0.11]	0.0664	[0.10]	0.0656	[0.10]	0.0631	[0.12]	-0.0314	[0.00]	-0.0301	[0.00]
	(0.040)		(0.040)		(0.040)		(0.040)		(0.010)		(0.010)	
Admission score	0.0072	[0.00]	0.0073	[0.00]	0.0072	[0.00]	0.0073	[0.00]	0.0015	[0.00]	0.0015	[0.00]
	(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
Share of female peers	0.3957	[0.18]	-0.3541	[0.24]	0.4484	[0.12]	-0.3096	[0.31]	0.0070	[0.92]	-0.0994	[0.19]
	(0.294)		(0.299)		(0.292)		(0.302)		(0.074)		(0.075)	
Share of SIPEE peers	-0.0390	[0.76]	-0.1446	[0.26]	-0.0459	[0.71]	-0.1803	[0.14]	0.0472	[0.16]	0.0288	[0.36]
	(0.126)		(0.128)		(0.125)		(0.123)		(0.033)		(0.032)	
Adm. Score of peers	-0.0017	[0.33]	-0.0072	[0.00]	-0.0011	[0.55]	-0.0073	[0.00]	-0.0013	[0.00]	-0.0001	[0.80]
	(0.002)		(0.000)		(0.002)		(0.000)		(0.000)		(0.000)	
Constant	-4.0684	[0.00]			-4.5185	[0.00]			-0.0625	[0.81]		
	(1.240)				(1.300)				(0.258)			
Observations	5,128		5,128		5,128		5,128		5,128		5,128	
Cohort fixed effects	No		Yes		No		Yes		No		Yes	

Note: Standard errors are reported in parentheses, P-values are reported in squared brackets.

If we now look at effects over career GPA, we again find a positive endogenous effect when we add cohort fixed effects, we also have gender and admission scores effects with the same sign, and yet again we only find a negative exogenous effect of the admission score of peers over career GPA when we add cohort fixed effects. Finally, if we look at the effects over passing rates, now we find a positive endogenous effect when we don't have fixed effects, there is now a positive female effect, a negative SIPEE student effect, again a positive admission score effect, and our only statistically significant exogenous peer effect is the negative effect of admission score of peers over passing rates when we don't have fixed effects.

The previous results are not robust at all, since adding cohort level fixed effects changes drastically the coefficient. This is not surprising, since this model is not identified, thus we should not care much about their results. Because of that, we are going to estimate again equation (3), but now we are going to use the generalized 2SLS technique described in Bramoulle et al. (2009) and Lee (2003), where we take advantage of the network structure to build an instrument for the endogenous effects. I implemented the routines of Bramoulle et al. (2009) for models of the same form of equation (3) that can include network fixed effects.

In Table 8 we observe the estimates obtained using the generalized 2SLS approach: if we look at the effects over the 1st semester GPA, I don't find evidence of endogenous peer effects, again I find positive female and admission score effects, and no evidence of exogenous peer effects. If we turn our attention the career GPA results, we now find a positive endogenous effect for the model without fixed effects, again positive female and admission score effects, and now being a SIPEE student increases the career GPA in almost half of a standard deviation in both models.

If we look at the exogenous peer effects, in the model without fixed effects we now find evidence of negative effects of the share of SIPEE peers and the admission scores of the peers over the career GPA of a student.

Finally, if we look at passing rates, we find a positive and statistically significant endogenous effect in both cases, but the magnitude of the coefficients are not reasonable: a coefficient of 1.16 implies that an increase in percentage point in the average passing rate of your peers makes you 1.16 percentage points more likely to pass your classes. An effect bigger than 1 in magnitude is not very common after controlling for so many variables, but it is still somehow plausible. However, a coefficient of almost 200 makes no sense. If we look at the independent variables, we find a positive effect of the female dummy, a negative SIPEE effect and a positive admission score effect. Regarding exogenous peer effects, I find evidence of a positive effect of the share of SIPEE peers over passing rates, and a negative effect of the admission score of peers.

Table 8: Peer effects Generalized 2SLS estimates

Variable	**1st semester GPA**				**Career GPA**				**Passing rate**			
Outcome of peers	1.3725	[0.17]	15.0583	[0.43]	0.8074	[0.09]	2.8500	[0.66]	1.1647	[0.00]	199.4451	[0.04]
	(0.990)		(18.980)		(0.475)		(6.410)		(0.233)		(97.577)	
Female	0.2057	[0.00]	0.1896	[0.00]	0.3594	[0.00]	0.3614	[0.00]	0.0188	[0.00]	0.0265	[0.00]
	(0.018)		(0.032)		(0.026)		(0.033)		(0.005)		(0.007)	
SIPEE student	0.0420	[0.32]	0.0675	[0.22]	0.4187	[0.00]	0.4183	[0.00]	-0.0338	[0.00]	-0.0706	[0.01]
	(0.042)		(0.055)		(0.062)		(0.085)		(0.011)		(0.026)	
Admission score	0.0070	[0.00]	0.0070	[0.00]	0.0107	[0.00]	0.0106	[0.00]	0.0014	[0.00]	0.0022	[0.00]
	(0.000)		(0.000)		(0.001)		(0.001)		(0.000)		(0.000)	
Share of female peers	-0.2044	[0.69]	-3.9096	[0.51]	-0.3777	[0.41]	-0.4082	[0.85]	-0.0471	[0.59]	0.2957	[0.18]
	(0.510)		(5.880)		(0.463)		(2.116)		(0.088)		(0.221)	
Share of SIPEE peers	0.0635	[0.64]	-0.9075	[0.57]	-0.5079	[0.01]	-0.5728	[0.75]	0.0732	[0.07]	3.7536	[0.00]
	(0.137)		(1.589)		(0.184)		(1.808)		(0.040)		(1.290)	
Adm. Score of peers	-0.0086	[0.11]	-0.0949	[0.48]	-0.0106	[0.01]	-0.0131	[0.82]	-0.0015	[0.00]	-0.2980	[0.05]
	(0.005)		(0.136)		(0.004)		(0.058)		(0.000)		(0.151)	
Constant	1.1351	[0.78]			-0.0152	[1.00]			-0.0862	[0.72]		
	(3.972)				(2.792)				(0.244)			
Observations	5,128		5,128		5,126		5,126		5,128		5,128	
Cohort fixed effects	No		Yes		No		Yes		No		Yes	

Note: Standard errors are reported in parentheses, P-values are reported in squared brackets.

Sadly, I don't think we can trust this much these results either. Even if I don't have a weak instruments test, the differences in magnitudes for the endogenous effects, and their huge standard errors are making me suspect that maybe I have a weak instruments problem. In future revisions I will work on developing a weak instruments test for this method.

Concluding remarks

In this paper I use network data from undergraduate students in Chile for a school that implemented an affirmative action program, and I try to study how these specially admitted students interact with their peers that got in the school through regular admission, and how their peer network (determined through random assignment to classes) can shape their outcomes and their interactions as well. At the student/class level I found that the class composition matters, with the number of SIPEE students in classroom increasing the grades of everyone, increasing the passing rates of other SIPEE students, but decreasing the passing rates of non-SIPEE students.

Regarding network characteristics, there is limited evidence of its importance, I only found that the number of peers in the opposite group to the one where a student belongs affects positively their career GPA and negatively their passing rate, while the clustering of the peers in the same group where a student belongs affects negatively their career grades and positively their passing rates. Models with peer effects so far have not yielded consistent results, but the generalized 2SLS estimates suggest the existence of positive endogenous peer effects, and a negative exogenous peer effect of the admission score of peers over outcomes.

The results here are only a starting point, since more complex models are needed here to try to shred some light about the existence of peer effects, classroom composition effects and network characteristics effects.

References

Ammermueller, A., & Pischke, J.-S. (2006). Peer effects in european primary schools: Evidence from pirls.

Bramoulle, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, *150*(1), 41–55.

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, *27*(3), 439–464.

De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, *2*(2), 241–75.

Hanushek, E. A., Kain, J. F., Markman, J. M., & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of applied econometrics*, *18*(5), 527–544.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation.

Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation: School reassignment and the structure of peer effects* (tech. rep.). Citeseer.

Kang, C. (2007). Classroom peer effects and academic achievement: Quasi-randomization evidence from south korea. *Journal of Urban Economics*, *61*(3), 458–495.

Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic perspectives*, *22*(3), 189–206.

Lee, L.-f. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*, *22*(4), 307–335.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), 531–542.

McEwan, P. J. (2003). Peer effects on student achievement: Evidence from chile. *Economics of education review*, *22*(2), 131–141.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, *116*(2), 681–704.

Whitmore, D. (2005). Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment. *American Economic Review*, *95*(2), 199–203.

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

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