Peer Effects in College Applications*

Katarína Borovičková Angela Crema Abby Holm Anusha Nath

October 2023

Abstract

This paper examines the effect of peers on high school students' decision of whether and where to apply for colleges. We use a rich administrative data from a large school district in Minnesota to identify peers based on detailed information about courses taken by students and random classroom assignments. We identify the effect of peers by using an instrumental variables strategy where we instrument the choices made by a student's direct peers by the decisions made by the peers of their peers. We find that peers have a significant positive effect on a student's probability of applying. Moreover, the average quality of one's peers' applications increases own application quality sizably and significantly. The effect is virtually zero for the lowest performing students, but it gets larger as own GPA raises. Peer effects seem to be at work within ability groups: high-performing students are mostly affected by their high-performing peers, whereas low-performing students respond more to the decisions of their low-performing peers.

JEL: I21, I23, I24

Keywords: Peer effects, college quality, college applications

*Borovičková: Federal Reserve Bank of Richmond, Crema: New York University, Holm: Hopkins Public Schools, Nath: Federal Reserve Bank of Minneapolis. We thank Meghana Bhaskar for excellent research assistance. The findings, conclusions, views, and opinions are solely those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Minneapolis, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

1 Introduction

There has been an increase in the college earnings premium in the last few decades and college quality has been identified to be one of the mechanisms that amplify the return to college for minority and disadvantaged students (Dale and Krueger, 2014). Despite the salience of this dimension to the debate on how to reduce inequality of opportunity, little is known on how students choose to apply at different institutions,[1] and how the choice of the institution (and its quality) affects their long run outcomes. In this paper we tackle these questions by focusing on one candidate factor that might enter high school students' decisions on where to apply for college: their classmates. While peer effects have been largely documented in the context of major choice (Sacerdote, 2001; De Giorgi et al., 2010; Feld and Zölitz, 2018; Zölitz and Feld, 2021), to the best of our knowledge this is the first analysis of peer effects on students' decision on where to apply for college. We answer this question using student-level administrative data from a large school district in Minnesota.

Literature on social interactions has to deal with two identification challenges: endogeneity and reflection. Endogeneity refers to a set of challenges that can arise when the effect of peers on an individual's outcome can be driven by unobservable factors that affect both the outcomes of the individual and of the peers. This includes the fact that individuals choose their own peer groups. The reflection problem is that one's behaviour affects everyone in the group, it is not possible to say whether ones' actions are their own or whether they are just a reflection of actions of someone else.

To address these concerns, we exploit De Giorgi et al. (2010)'s approach which relies on partially overlapping groups of peers which are such that peer groups of two peers do not perfectly

[1] There is an ample literature on the determinants of major choice: see Patnaik and Zafar (2021) for a comprehensive review.

overlap. De Giorgi et al. (2010) show that it is possible to use partially overlapping groups to address both the endogeneity and the reflection problem.

The reflection problem is a type of simultaneous equation model where the dependent variable (in our case outcome of a student) is a function of another dependent variable (outcomes of her peers). One way to achieve identification nn a simultaneous equations model is to impose parameter restrictions and exclusion restriction. Intuitively, excluded peers play the role of exclusion restriction since they affect outcomes of the peers, but not of the student i directly.

We address the endogeneity problem due to group-specific factors using exogenous characteristics of the excluded peers as an instrument. These are correlated with peers' outcomes due to social interactions, but are uncorrelated with shocks common to the group because they are excluded.

We use data from a large school district in Minnesota. In the last year of high school, students choose which classes they want to attend. Most classes are large and hence split into several sections at random. Our data contains information on the composition of students in each section. We define peer groups of student i as consisting of all classmates of i across sections that student i attends. Since sections are created randomly, this process naturally forms partially overlapping peer groups. We rely on this information for identification of the peer effects: because peer groups in high school are not perfectly overlapping, we instrument student i's direct peers (classmates) with i's excluded peers, that is, students who are in class with at least one of i's peers, but not with i directly.

The excluded peers serve two goals. They play the role of exclusion restriction to address the reflection problem which is a special case of the simultaneity problem. Moreover, the exogenous characteristics of excluded peers serves as an instrument to deal with endogeneity generated by common group effects. Due to the random assignment of students into sessions, we do not have to deal with the endogeneity issue arising from students choosing their own peer group.

We find that one additional peer applying for college increases the own probability of applying by around 0.5 percentage points. Moreover, an increase in the average quality of applications by peers increases own application quality sizably and significantly. The effect is virtually zero for the lowest performing students, but it gets larger as own GPA raises. Peer effects seem to be at work within ability groups: high-performing students are mostly affected by their high-performing peers whereas low-performing students respond more to the decisions of their low-performing peers.

We contribute to a broad literature that studies the impact of peers on students' outcomes in secondary schools and colleges. Our paper is closest to De Giorgi et al. (2010), both in terms of question and the methodology. They study peer effects in the choice of college major and find strong evidence that an individual is more likely to choose a major when many of her peers make the same choice. They also find that peers can divert students from majors in which they have a relative ability advantage, with adverse consequences on academic performance, entry wages, and job satisfaction.

Our results showing that females and White students are less responsive to their peers adds to the literature that examines how peer effects in schools depend on gender and race. Hoxby (2000) exploits within school variation in the gender and racial composition of classrooms in adjacent years in schools in Texas. She finds that students are affected by achievements of their peers. While she does not find evidence of non-linearities in peer effects, she finds that peer effects are stronger within racial groups than between groups. She further finds that a larger share of females in class improves outcomes for both male and female peers. Lavy and Schlosser (2011), reach a similar conclusion. They find that an increase in the proportion of girls improves boys' and girls' cognitive outcomes. These academic gains are mediated through lower levels of classroom disruption and violence, rather than the higher achievements themselves. On the contrary, Black et al. (2013) find that higher proportion of females in the ninth grade decreases educational attainment and probability of selecting an academic track. At the same time, it leads to lower teenage birth rates and higher earnings for women. Anelli and Peri (2019) find that male students attending a high school class with more than eighty percent male classmates have a higher probability of choosing a prevalent male major, but find no long-lasting effects on graduation and labor market outcomes. Females are not affected by the gender class composition.

The evidence on peer effects on academic outcomes is mixed. While Whitmore (2005) and Hanushek et al. (2003) find evidence of positive peer effects on academic outcomes, several other studies find small or short-lived peer effects (Burke and Sass (2013), Angrist and Lang (2004) or Imberman et al. (2012)). We do not study academic outcomes per se, but explore how peer effects for college applications depends on the academic outcomes of students.

Our paper examines peer effects separately for students with different levels of GPA and finds heterogeneous effects. This is in line with a related literature which argues that linear in means model of peer effects is not adequate as peer effects are non-linear. Hoxby and Weingarth (2005) exploit changes to school assignment policies in Wake County, in North Carolina. They allow peer effects to depend on student's own position in test score distribution and type of peers being expanded. They find that high (low) achievers benefit from adding high (low) achievers. Similar results found by Gibbons and Telhaj (2016) for United Kingdom. Burke and Sass (2013) find that for low-achieving students, having very-high-aptitude peers appears worse than having peers of average ability. Lavy et al. (2012) use data for secondary school in United Kingdrom to find that everyone is harmed by low-achieving peers, while only high-achieving girls benefit from high achievers.

Several studies find modest peer effect on achievement in colleges. Sacerdote (2001) finds model impact of peers on students' academic performance, while Zimmerman (2003) shows that students are harmed by bottom 15% of roommates. Stinebrickner and Stinebrickner (2006), Fletcher and Tienda (2010) find that college students benefit from peers they attended high school with. Our paper is related to these as we highlight the role of peers in the choice of applying to college itself.

The rest of this paper is structured as follows: Section 2 describes the data; Section 3.1 defines own and excluded peers and illustrates the identification approach; Section 3.2 describes the results; Section 4 concludes.

2 Data

Our analysis is based on administrative data from a large school district in Minnesota. We complement this dataset with information from Naviance, a system that students use to send applications to colleges, and data on college quality collected by Integrated Postsecundary Education Data System (IPEDS). Our analysis uses full cohort of students in this school district which graduated from high school between 2009 and 2022.

We use information on students enrolled in Grade 12 of the public high school in this school district. Students in this cohort take several mandatory classes and several optional classes of their choice. Most of the classes are large and are split into several sections at random. Importantly, every section of the class has a unique identifier which allows us to identify students who take the same class in the same section. We also observe students' grades in each class taken. The average size of a section is 21 students.

We classify classes into instructional and non-instructional using course titles. We adopt a restrictive definition of an instructional course, one that only includes academic and non-

Table 1: Descriptive statistics of students in our sample

Variable	Mean	SD	Min	Max	N
no. of applications	5.09	4.29	1	56	6,546
1 = female	0.51	-	0	1	6,546
1 = white	0.69	-	0	1	6,546
1 = lowincome	0.32	-	0	1	6,546
GPA	3.23	0.64	0	4	6,546

N in the last column refers to the number of students.

vocational courses. Examples of courses in our sample that we label as non-instructional are Ceramics, Culinary, Drawing and Painting, Guitar, Jewel and Metal-smithing, Sculpture, and Video Production. We use this distinction between types of classes in the definition of the peer group.

Students send applications to colleges in Grade 12. In our sample, 83% of students sent at least one application. Our analysis of peer effects focuses on these students. More than a half of our sample sent four applications or less, with 21% of the sample sending only one application. Table 1 shows the distribution of applications in our sample, along with other descriptive statistics on gender, race, and socio-economic composition of students.

The data set from the school district is combined with information from Naviance. Due to privacy reasons, we do not observe college names, only a unique identifier assigned to them by the school district. For each college, we observe whether it is 2-year or a 4-year college and the state in which it is located.

We further complement the datasets described above with information on college quality collected by Integrated Postsecundary Education Data System (IPEDS) which is publicly available.[2] The IPEDS information is available for around 1,000 colleges in the U.S. IPEDS provides several statistics on the incoming cohort. We use 75^{th} percentile of the ACT math score in 2018

[2]The school district merged the Naviance data with IPEDS' quality measure using college names before anonymizing the college names.

Table 2: Distribution of colleges across quality categories

ACT math score	16–22	23	24	25	26	27	28	29	30	31	32–36
U.S. college	0.18	0.10	0.11	0.14	0.16	0.10	0.06	0.04	0.03	0.02	0.06
in sample applications	0.01	0.03	0.05	0.06	0.06	0.12	0.12	0.05	0.13	0.06	0.14

Notes: The first row shows the distribution of colleges across quality categories in IPEDS. The second row shows the distribution of applications in our sample across quality categories, conditional on the college being ranked in IPEDS.

as a measure of college quality.

In our sample, students sent a total of 33,403 college applications. Out of these, 222 were sent to schools outside of the U.S. and 405 to institutions which are not ranked in IPEDS. The ACT math test is on a scale 1–36. In our sample, 9% of applications were sent to colleges with ACT math score below 25, 14% to colleges above 32 points. Most applications, 42%, were sent to colleges with the ACT math score between 27–30 which in the ranking accounts for about 23% of colleges. Overall, students tend to send applications to better colleges than what corresponds to a "random" sample of colleges, as shown in Table 2.

3 Peer Effects

3.1 Empirical Strategy

We start by defining two objects of interest for each student, the peer group and the excluded peer group. For student i, we define peer group G_i as the set of all students j who take at least one class in grade 12 with student i. We further define the peers of peers group for student i, $\tilde{G}_i$, to include students k such that $k \in G_j$ for $j \in G_i$ but $k \notin G_i$. In words, the set $\tilde{G}_i$ is the set of students who are peers with at least one peer of i, but they are not peers of i directly.

As is standard in the literature, the key assumption is that these classes capture the network that students interact in. Moreover, it is assumed that students who do not take any classes together do not interact. We focus on grade 12 as we expect this grade to have the strongest influence on college-related decisions. We first consider all classes that student i takes to the define the peer group. However, it is possible that some classes do not offer much room for interactions among students and hence these create potentially weaker links among themselves. To address this concern, we consider two alternative definitions of peer groups, one based on only instructional classes, and the other based on only non-instructional classes.

We seek to estimate the linear-in-means model

$$y_i = \alpha + \beta E(y|G_i) + \gamma E(\boldsymbol{x}|G_i) + \delta \boldsymbol{x}_i + u_i, \tag{1}$$

where y_i is chosen college, $\boldsymbol{x}_i$ is a set of individual characteristics, $E(\boldsymbol{x}|G_i)$ is the average of individual characteristics of i's peers. The coefficient of interest is β which measures the endogenous effect of peers' outcomes on outcomes of student i. The coefficient γ is the so-called exogenous effect.

In a standard peer group setting, if students i and j are classmates, then their peer groups coincide and $G_i = G_j$. In this case, it is not possible to separately estimate β and γ and so the total effect combines the endogenous and exogenous effects together. This is explained in De Giorgi et al. (2010), but for clarity we show the argument here as well. Consider a student j who is peers with i, $i \in G_j$, and take average of (1) across all $i \in G_j$. Assuming that $E(u_i|G_j) = 0$, we have

$$E(y_i|G_j) = \alpha + \beta E(E(y|G_i)|G_j) + \gamma E(E(\boldsymbol{x}|G_i)|G_j) + \delta E(\boldsymbol{x}_i|G_j). \tag{2}$$

If $G_i = G_j$, then $E(E(y|G_i)|G_j) = E(y|G_i)$ and the resulting equation is

$$E(y_i|G_j) = \frac{\alpha}{1-\beta} + \frac{\beta+\gamma}{1-\beta} E(\boldsymbol{x}|G_j), \tag{3}$$

and so β and γ are not separately identified.

The insight of De Giorgi et al. (2010) is that, in case of individual-specific peer groups, students i and j do not have perfectly overlapping peer groups, and hence $G_i \neq G_j$. In this case, the iterated expectation in (2) does not collapse, and we have variation in $E(y|G_i)$.

In deriving (2), we assumed that $E(u_i|G_j) = 0$ for $in \in G_j$, or in words, that there are no correlated group-specific effects. This is unlikely to be the case in class setting where students in a class are taught by the same teacher. Following the argument in De Giorgi et al. (2010), the non-overlapping peer groups help us address this problem as well. The excluded peers $\tilde{G}_i$ can be used as an instrument for the endogenous variable because they are not affected by group-specific shocks of student i (they are not classmates) but are correlated with the mean outcome of peers of i due to social interactions.

Finally, another typical challenge is endogeneity in forming peer's group. However, this problem does not arise in our setting because the sections are formed randomly. That is, conditional on the choice of class, students are assigned to sections randomly via an algorithm that the administration of the district has been using for several years now.[3]

We are interested in understanding the peer effects in submitting application to a particular college. We thus estimate

$$y_i^m = \alpha + \beta z_i^m + \gamma E(\boldsymbol{x}|G_i) + \delta \boldsymbol{x}_i + u_i, \tag{4}$$

where y_i^m equals 1 if student i applies to college m, z_i^m is the share of peers from G_i who applied to college m, and $\boldsymbol{x}$ is the vector of exogenous characteristics. We use the share of excluded peers of i who applied to college m as an instrument.

[3]Once courses have started, students can ask to be transferred to a different section by getting an appointment with their counselors, as long as this is compatible with the rest of their school schedule. The administrators of the classroom formation process have confirmed to us that this happens, although rarely.

Table 3: Peer effects in the choice of applying for colleges

VARIABLES	All courses (1)	Instructional (2)	Non-instructional (3)
% Peers Applying	0.929^{***}	0.928^{***}	0.994^{***}
	(0.0099)	(0.0096)	(0.0104)
College quality	Y	Y	Y
Student fixed effects	Y	Y	Y
Obs. (student x college)	2,725,457	2,651,172	2,724,109
N. students	6,546	6,367	6,543
N. colleges	1,333	1,333	1,333

Notes: Results from 2SLS regressions at the student-by-college level. Graduating cohorts 2009–2022 are included in the sample. The dependent variable takes value 1 if there is an application on record for the corresponding student-college pair, zero otherwise. The main explanatory variable is the fraction of own peers applying for a given college and is instrumented with the fraction of excluded peers who apply for that college, see Section 3.1 for details on the definition of own and excluded peers. Courses are classified into instructional and non-instructional based on the course titles. Standard errors are clustered at the student level.

3.2 Results

Table 3 shows the results obtained from regressions at the student-by-college level where the dependent variable is a dummy taking value 1 if there is an application on file for the corresponding student-college pair, 0 otherwise. The explanatory variable is the fraction of own peers applying for that same college and is instrumented by the share of excluded peers submitting an application to that college. All regressions include the measure of college quality described in Section 2 and student fixed effects, which absorb all the time-invariant characteristics that are specific to each student. Column 1 defines peers, both own and excluded, based on interactions in any school course, whereas columns 2 and 3 restrict to instructional and non-instructional courses, respectively.

As the share of peers that apply to a certain college increases from 0 to 100%, the probability that a student also applies to that college increases by more than 90 percentage points on average (Column (1)). In other terms, since the average size of the pool of own peers in our sample is approximately 200, one additional peer applying for a certain college increases the own probability of applying by around 0.5 percentage points. This estimate is small in magnitude compared to De Giorgi et al. (2010) finding that one additional peer choosing economics as a major increases the likelihood of opting for economics by 7.4 percentage points. Besides the difference between our research question and theirs, this is expected, as the average size of the peer group in their context is much smaller (around 18), which suggests that their larger marginal effect is probably due to stronger peer-to-peer interactions.

Interestingly, the magnitude and precision of the estimate reported in Column (1) of Table 3 do not vary as peers are defined within different types of courses. If anything, peer effects seem stronger within non-instructional courses, although the difference between the coefficients in columns (2) and (3) is not statistically significant. This finding is consistent with the interpretation that non-academic interactions can affect academic choices.

Table 4 explores how the peer effect documented in Table 3 varies across students of different genders, ethnic and socio-economic backgrounds, and levels of school performance. While females and males seem to be equally responsive to their peers' behavior (column (2)), students who are non-white (column (3)), or have a high GPA (column (5)) are more responsive to their peers' application decision. Column (5) shows that when we include all interaction terms in the same regression model, the interaction with GPA remains statistically and economically high.

As expected, the result that students "follow" their peers maps directly into the quality of the institutions that students end up applying for. Table 5 replicates the analysis in Table 4, with the differences that the regressions are at the student (not student-college) level, the dependent variable is the average quality of the colleges for which the student applies, and the explanatory variable and the instrument are the average quality of own and excluded peers' applications. The quality effect is large. Column (6), our preferred specification, suggests that as the average

Table 4: Individual heterogeneity in peer effects in application choices

VARIABLES	All courses (1)	All courses (2)	All courses (3)	All courses (4)	All courses (5)
% Peers Applying	0.929***	0.920***	0.866***	0.697***	0.744***
	(0.0099)	(0.0144)	(0.0193)	(0.0153)	(0.0223)
% Peers Applying*		0.017			−0.044*
*Female(own)		(0.0199)			(0.0196)
% Peers Applying*			0.091***		−0.054*
*White(own)			(0.0225)		(0.0228)
% Peers Applying*				0.407***	0.429***
*GPA(own)				(0.0197)	(0.0208)
College quality	Y	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y	Y
Obs. (student x college)	2,725,457	2,725,457	2,725,457	2,725,457	2,725,457
N. students	6,546	6,546	6,546	6,546	6,546
N. colleges	1,333	1,333	1,333	1,333	1,333

Notes: Results from 2SLS regressions at the student-by-college level. Graduating cohorts 2009–2022 are included in the sample. The dependent variable takes value 1 if there is an application on record for the corresponding student-college pair, zero otherwise. The main explanatory variable is the fraction of own peers applying for a given college and is instrumented with the fraction of excluded peers who apply for that college: see Section 3.1 for details on the definition of own and excluded peers. GPA is defined on a scale from 0 to 4, with larger values corresponding to a better average school performance. Standard errors are clustered at the student level.

quality of the peers' application increase by 1 point,[4] the average quality of own applications increases by more than 2 points for students with a GPA below the median, which is equivalent to an increase in average application quality for these sub-sample of students larger than 50%. Even though it is large, this effect is concentrated among students who are relatively close to the median GPA, as the lowest tail of the performance distribution seems unresponsive to peers' choices (see Table 4). The quality effect on students with above-median GPA is slightly smaller, as indicated by the negative and statistically significant interaction term, and corresponds to an

[4]See Section 2 for details on how our quality scale is defined.

increase in quality of almost 30% within this sub-sample of students.

The results above suggest that students are influenced by their peers in their college application decisions, and a higher (own) GPA correlates with a larger peer effect. But how do peers of different performance types affect students of a certain type? Table 6 shows the results of regressions similar to those estimated in Table 5, with the difference that we distinguish between peers with a GPA above and below the median.[5] Both high- and low-achieving peers seem to influence students' choices, and the magnitude of the two peer effects is not statistically distinguishable (column (1)). However, when we interact peers' quality with an indicator for whether the student herself has a high or low GPA, some interesting heterogeneity emerges. The effect of high-GPA peers appears stronger on high-GPA students, as indicated by the positive (yet imprecisely estimated) interaction term, whereas low-GPA peers have a larger impact on low-GPA students (negative interaction term). In other words, peer effects seem to be at work with ability groups, where GPA is used as a proxy for ability.

4 Conclusion

In this paper we study school peer effects on students' decisions of where to apply for college. To do so, we exploit non-fully overlapping groups of peers (classmates) for identification of the peer effect, alongside rich administrative data at the student level. We find that an increase in the average quality of one's peers' applications increases own application quality sizably and significantly. The effect is virtually zero for the lowest performing students, but it gets larger as own GPA raises. Peer effects seem to be at work within ability groups. High-performing students are mostly affected by their high-performing peers, whereas low-performing students respond more to the decisions of their low-performing peers.

[5] In this specification, our instruments are the average application quality of the high-GPA excluded peers and the low-GPA excluded peers (alongside interactions).

Table 5: Peer effects in application quality

VARIABLES	All courses (1)	All courses (2)	All courses (3)	All courses (4)	All courses (5)	All courses (6)
Average quality of peers' applications	1.971***	2.199***	1.713***	2.109***	3.333***	3.885***
	(0.0921)	(0.1554)	(0.3254)	(0.1857)	(0.4299)	(0.6684)
Female (own)	−0.553***	2.452	−0.551***	−0.554***	−0.478***	2.641
	(0.0716)	(1.813)	(0.0716)	(0.0716)	(0.0756)	(2.0198)
White (own)	0.027	0.026	−2.266	0.067	0.097	2.410
	(0.0972)	(0.0970)	(2.8155)	(0.1081)	(0.1135)	(3.1149)
Economically Disadvantaged (own)	−0.674***	−0.683***	−0.700***	2.539	−0.673***	1.699
	(0.0989)	(0.0989)	(0.1045)	(3.6129)	(0.1278)	(4.219)
GPA (own)	2.106***	2.098***	2.130***	2.145***	15.326**	14.923**
	(0.1162)	(0.1146)	(0.1191)	(0.1207)	(5.4593)	(5.627)
Peers' app quality* *Female (own)		−0.462				−0.481
		(0.2772)				(0.3089)
Peers' app quality* *White (own)			0.362			−0.360
			(0.4424)			(0.4883)
Peers' app quality* *Ec. disadv. (own)				−0.508		−0.373
				(0.5698)		(0.6615)
Peers' app quality* *High GPA (own)					−2.054*	−1.988*
					(0.8349)	(0.8610)
Graduating year (own)	Y	Y	Y	Y	Y	Y
Obs. (students)	6,546	6,546	6,546	6,546	6,546	6,546
Avg dep. var.	6.405	6.405	6.405	6.405	6.405	6.405

Notes: Results from 2SLS regressions at the student level. Graduating cohorts 2009–2022 are included in the sample. The dependent variable is the average quality of the colleges for which a student applies. College quality is defined into 16 categories based on the 75th percentile of the admission Math SAT score: Larger values of the variable correspond to higher scores, with 0 denoting the lowest score bin and 16 the highest score bin. The main explanatory variable is the average quality of the colleges where own peers apply and is instrumented with the average quality of the colleges where excluded peers apply: See Section 3.1 for details on the definition of own and excluded peers. High and low GPA are defined relative to the sample median. GPA is defined on a scale from 0 to 4, with larger values corresponding to a better average school performance.

Table 6: Peer effects in application quality by peers' GPA

VARIABLES	All courses (1)	All courses (2)
High-GPA peers' app quality	1.561***	1.388***
	(0.1403)	(0.1385)
High-GPA peers' app quality* * High-GPA (own)		0.123
		(0.1309)
Low-GPA peers' app quality	1.268***	1.369***
	(0.1098)	(0.1457)
Low-GPA peers' app quality* *High-GPA (own)		−0.089
		(0.2437)
GPA (own)	2.558***	2.183***
	(0.0986)	(0.1387)
Female (own)	Y	Y
White (own)	Y	Y
Economically disadvantaged (own)	Y	Y
Graduating year	Y	Y
(Obs.) N. students	6,537	6,537
Average dep. var.	6.389	6.389

Notes: Results from 2SLS regressions at the student level. Graduating cohorts 2009–2022 are included in the sample. The dependent variable is the average quality of the colleges for which a student applies. College quality is defined into 16 categories based on the 75th percentile of the admission Math SAT score: Larger values of the variable correspond to higher scores, with 0 denoting the lowest score bin and 16 the highest score bin. The main explanatory variable is the average quality of the colleges where own peers apply and is instrumented with the average quality of the colleges where excluded peers apply: See Section 3.1 for details on the definition of own and excluded peers. Peers are divided between high and low GPA, which are defined relative to the sample median. GPA is defined on a scale from 0 to 4, with larger values corresponding to a better average school performance.

From a policy perspective, these results raise the concern that schools may not be able to improve the educational path that the most disadvantaged, lowest-performing students select into by exposing them to classmates who apply for better colleges. Also, the fact that peer effects seem to be at play within ability groups suggests that, from the point of view of efficiency, the classroom formation decision implies trade-offs. For instance, in a scenario where it is desirable to have all students apply for high-quality colleges (to increase their pre-graduation effort, for

Table 7: Individual heterogeneity in peer effects in attending college

VARIABLES	All courses (1)	All courses (2)	All courses (3)	All courses (4)	All courses (5)
% Peers Applying	0.914***	0.924***	0.929***	0.740***	0.849***
	(0.0186)	(0.0263)	(0.0334)	(0.0241)	(0.0374)
% Peers Applying*		−0.020			−0.074
*Female(own)		(0.0371)			(0.0376)
% Peers Applying*			0.021		−0.146**
*White(own)			(0.0401)		(0.0427)
% Peers Applying*				0.310***	0.362***
*GPA(own)				(0.0361)	(0.0395)
College quality	Y	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y	Y
Obs. (student x college)	2,725,457	2,725,457	2,725,457	2,725,457	2,725,457
N. students	6,546	6,546	6,546	6,546	6,546
N. colleges	1,333	1,333	1,333	1,333	1,333

Notes: Results from 2SLS regressions at the student-by-college level. Graduating cohorts 2009–2022 are included in the sample. The dependent variable takes value 1 if a student chooses to attend a college for the corresponding student-college pair, zero otherwise. The main explanatory variable is the fraction of own peers applying for a given college and is instrumented with the fraction of excluded peers who apply for that college: see Section 3.1 for details on the definition of own and excluded peers. GPA is defined on a scale from 0 to 4, with larger values corresponding to a better average school performance. Standard errors are clustered at the student level.

example), schools should compare the imitative/motivational gain for high-achieving students when grouped within the same class to the price that low-achieving students pay from negatively influencing one another. This project will evolve in this direction by trying to understand the mechanisms underlying the peer effects herein documented and studying how the decision of where to apply for college relates to pre-graduation effort, the choice of major, aggregate learning and skill accumulation, and the gender and racial gap in access to high-quality higher education.

References

Anelli, Massimo and Giovanni Peri (2019), “The effects of high school peers’ gender on college major, college performance and income.” *The Economic Journal*, 129, 553–602.

Angrist, Joshua D and Kevin Lang (2004), “Does school integration generate peer effects? evidence from boston’s metco program.” *American Economic Review*, 94, 1613–1634.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes (2013), “Under pressure? the effect of peers on outcomes of young adults.” *Journal of Labor Economics*, 31, 119–153.

Burke, Mary A and Tim R Sass (2013), “Classroom peer effects and student achievement.” *Journal of Labor Economics*, 31, 51–82.

Dale, Stacy B. and Alan B. Krueger (2014), “Estimating the effects of college characteristics over the career using administrative earnings data.” *Journal of human resources*, 49, 323–358.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli (2010), “Identification of social interactions through partially overlapping peer groups.” *American Economic Journal: Applied Economics*, 2, 241–275.

Feld, Jan and Ulf Zölitz (2018), “Peers from venus and mars–higher-achieving men foster gender gaps in major choice and labor market outcomes.”

Fletcher, Jason and Marta Tienda (2010), “Race and ethnic differences in college achievement: Does high school attended matter?” *The Annals of the American Academy of Political and Social Science*, 627, 144–166.

Gibbons, Stephen and Shqiponja Telhaj (2016), “Peer effects: Evidence from secondary school transition in england.” *Oxford Bulletin of Economics and Statistics*, 78, 548–575.

Hanushek, Eric A, John F Kain, Jacob M Markman, and Steven G Rivkin (2003), “Does peer ability affect student achievement?” *Journal of applied econometrics*, 18, 527–544.

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

Hoxby, Caroline M and Gretchen Weingarth (2005), “Taking race out of the equation: School reassignment and the structure of peer effects.” Technical report, Citeseer.

Imberman, Scott A, Adriana D Kugler, and Bruce I Sacerdote (2012), “Katrina’s children: Evidence on the structure of peer effects from hurricane evacuees.” *American Economic Review*, 102, 2048–2082.

Lavy, Victor and Analia Schlosser (2011), “Mechanisms and impacts of gender peer effects at school.” *American Economic Journal: Applied Economics*, 3, 1–33.

Lavy, Victor, Olmo Silva, and Felix Weinhardt (2012), "The good, the bad, and the average: Evidence on ability peer effects in schools." *Journal of Labor Economics*, 30, 367–414.

Patnaik, Matthew Wiswall, Arpita and Basit Zafar (2021), "College majors 1." *Routledge*.

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

Stinebrickner, Ralph and Todd R Stinebrickner (2006), "What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds." *Journal of public Economics*, 90, 1435–1454.

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

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

Zölitz, Ulf and Jan Feld (2021), "The effect of peer gender on major choice in business school." *Management Science*, 67, 6963–6979.