Study-Group Diversity and College Academic Outcomes: Experimental Evidence from the Beca18 Social Inclusion Program in Peru[1]

M. Agurto[2 a, b], H. Fiestas [a], W. Nuñez [a], V. Quevedo [a, c], S. Vegas [a]

[an] Universidad de Piura

[b] Lima School of Economics

[c] Virginia Tech

Abstract

Beca18 is a social inclusion program in higher education implemented by the Peruvian government. It provides full scholarships to poor public high school graduates who have been admitted into an eligible higher education institution. We exploit random assignment of freshman students to small study-groups at a private university in Peru to evaluate how the interactions between Beca18 fellows and the relatively wealthier non-Beca18 students who predominantly attend this institution affect academic outcomes. We find that assignment to mixed study-groups, composed of Beca18 and non-Beca18 students, on average improves performance on individual weekly/biweekly quizzes. In the case of Beca18 fellows this effect is increasing in a pre-college academic performance indicator; that is, relatively high ability Beca18 fellows benefit the most. Interestingly, the opposite holds for non-Beca18 students. Our evidence points to two underlying mechanisms behind these results: (i) peers' academic ability and (ii) peer's attitudes towards effort and cooperation. In the case of Beca18 students assigned to mixed groups, they seem to benefit from the relatively higher academic ability of their non-Beca18 peers. In the case of non-Beca18 students, they appear to benefit from the higher effort and cooperation displayed by Beca18 fellows. The benefits of mixed study-group assignment also seem to be stronger among non-Beca18 students with relatively low household incomes; as well as in mixed groups containing the same number of Beca18 and non-Beca18 individuals.

[1] This research has been generously funded by the Government of Japan and supported by the Global Development Network (GDN) through the Global Development Awards Competition, Japanese Award for Outstanding Research on Development 2016. We thank German Vega and Paul J Corcuera for their excellent research assistance and Mario Palacios for his extraordinary logistic support. We thank participants at the Canadian Association Meetings 2017, Peruvian Economic Association Conference 2017, Virginia Tech Economics Department Seminar, UDEP Engineering Faculty Innovation Seminars, UDEP Economics Department Seminar, 8th Bolivian Economic Development Conference 2017, LAWEBESS 2017 at Universidad del Valle, and GRADE Seminar for their comments on previous versions of this paper. We specially thank Mario Picon for his valuable comments and guidance. The usual disclaimer applies.

[2] Corresponding author: marcos.agurto@udep.pe +51 73 284500 ext. 3134. Departamento de Economía - Universidad de Piura. Oficina 218 – Edificio E. Av. Ramón Mugica 131 – Urb. San Eduardo – Piura, Peru.

1. Introduction

As a way to promote social mobility and alleviate social inequalities, in recent years several governments in the developing world have increased their efforts to improve the access of socioeconomically deprived youth individuals into top-quality higher education institutions (Beyer et al., 2015; Melguizo et al., 2016). Concrete examples of such efforts are the public funded programs SerPiloPaga in Colombia, PoliticadeCuotas in Ecuador and Beca18 in Peru. These programs provide comprehensive financial aid to poor high school graduates admitted into eligible, mostly private, elite universities; increasing as a result socioeconomic diversity at highly selective post-secondary institutions. While a few recent studies have evaluated these programs' impacts on postsecondary enrollment and dropout rates (Melguizo et al., 2016; Londoño-Velez, 2017); there is scarce evidence on how the academic interactions between the wealthy students who predominately attend elite universities and the poor students targeted by these programs influence key individual outcomes, such as academic performance, social preferences and behavior.

In this paper we exploit experimental evidence related to the Peruvian program Beca18 to estimate the academic peer effects associated with students' interactions in a socioeconomically diverse college environment. Implemented since 2012, Beca18 is a government funded program which provides full scholarships to public high school graduates who live in poverty according to the Peruvian Focalization Index (SISFOH by its Spanish initials), obtained a relatively high academic achievement during their upper high school years (measured in terms of high school GPA), and have been admitted into an eligible, high-quality, private university.

More specifically, we study a Randomized Control Trial (RCT) at a private university in Peru, Universidad de Piura (UDEP), in which 568 freshman students in the 2016 cohort were randomly assigned to small study-groups. While some students were assigned to (mixed) study-groups containing both Beca18 fellows and relatively wealthier non-Beca18 students (mostly middle and upper-middle class individuals[3]); others were assigned to (unmixed) groups containing only Beca18 or only non-Beca18 individuals. Students meet in these study-groups during weekly tutorial labs to review the material covered in lectures and prepare in this way for their weekly/biweekly individual quizzes[4]. Tutorial labs are however not held during the midterm and final exams periods, lasting 10 to 14 days each. It is important to highlight that students are not required to work or study together in their randomly assigned study-groups outside tutorial labs, or during midterm and final exams. Moreover, tutorial labs usually end one to two weeks before the final exams begin.

[3] Less than 3% of the non-Beca18 students in our study sample qualify as monetarily poor.

[4] These quizzes ("practicas calificadas" in Spanish) are common in the Peruvian college system. In the case of UDEP, they account for 30% to 40% of the final course grade.

In the context of the Beca18 program, it is not ex-ante obvious how the assignment of Beca18 and non-Beca18 students to the same study-group will affect their academic outcomes. Consider in first place effects related to peers' academic ability. Several studies in the literature suggest that relatively low ability students can significantly benefit from interacting with high ability peers (Stinebrickner and Stinebrickner, 2006; Lyle, 2007; Carrel, 2009; Brunello et al, 2010). Nevertheless, as Peru lacks a national high school exit test or universal college admission exam, directly assessing the relative pre-college differences in academic ability between Beca18 and non-Beca18 students and hypothesizing about which group will likely benefit from being assigned higher ability peers is problematic. Furthermore, we cannot rely on high school GPA comparisons, given the critical academic differences between public high schools in socioeconomically deprived areas attended by Beca18 fellows and the private schools attended by high-income individuals. Moreover, anecdotal evidence suggests that several eligible universities applied differentiated admission mechanisms and criteria to Beca18 and non-Beca18 students, which complicates using admission test scores to academically compare Beca18 and non-Beca18 students.

In the specific case of UDEP, all Beca18 fellows attend a six months remedial program before the first term, which aims to address the potential gaps in high school preparation among Beca18 students. On the other hand, admitted non-Beca18 students attend a separate highly intensive four weeks summer program, which main objective is to assess their high school academic background. We take the GPA obtained by Beca18 students in their remedial program as an indicator of their pre-college academic ability and use it to academically rank Beca18 fellows. We proceed similarly with non-Beca18 students but use in this case the GPA obtained in the four weeks summer program. However, given the different nature and objectives of the remedial and summer programs, the related Beca18 and non-Beca18 pre-college indicators are not comparable; and cannot be used to assess the relative differences in pre-college academic ability across Beca18 and non-Beca18 individuals.

Besides, even if we could provide direct, ex-ante, evidence on the relative pre-college academic differences among Beca18 and non-Beca18 students, socioeconomic segregation within study-groups may still take place (i.e. group members may decide to assign students from the same socioeconomic status to the same tasks), limiting the scope of the benefits linked to academic peer effects. Moreover, peer effects associated to relative differences in academic ability may not be the only type of peer effects affecting individual outcomes in the Beca18 program context. Several instructors we interviewed at UDEP indicated that Beca18 fellows are highly motivated, persistent, hardworking and collaborative. Such personality traits, as has been suggested by recent works in the

peer effects literature (i.e. Golsteyn et al., 2017), may influence the attitudes, motivation, and as a result the academic performance of their relatively wealthier non-Beca18 peers[5].

In this paper we explore how mixed study-group assignment affects academic performance on two types of individual evaluations: i) weekly/biweekly quizzes ("practicas calificadas" in Spanish) and ii) midterm and final exams. In the case of weekly/biweekly quizzes we find that, on average, freshman students in mixed study-groups perform better than those in unmixed ones. We also test for heterogenous effects related to pre-college academic performance by estimating separate regressions for Beca18 and non-Beca18 students and including an interaction term between the treatment dummy and the group-specific pre-college academic indicator discussed above. Our results indicate that among Beca18 fellows, the treatment effect on weekly/biweekly quizzes is increasing in the Beca18 pre-college academic performance indicator. That is, relatively high academic ability Beca18 fellows benefit the most from mixed group assignment. Among non-Beca18 individuals the opposite pattern is observed. Our evidence also suggests that some students are negatively affected in terms of their weekly/biweekly quizzes performance. These are the non-Beca18 students in the top 15% and the Beca18 students in the bottom 20% of their group specific pre-college academic performance distribution.

In the case of midterm and final exams, the estimated average treatment effect is relatively close to zero for both non-Beca18 and Beca18 students, as well as not statistically significant. We also find evidence of heterogenous effects related to precollege academic ability, but only among non-Beca18 students. More specifically, as in the case of weekly/biweekly quizzes we find that relatively low ability non-Beca18 student benefit the most, while some relatively high ability students are negatively affected. Interestingly, the proportion of non-Beca18 students negatively affected is higher in this case than the observed during weekly/biweekly quizzes. We hypothesize that the differences in treatment effects on weekly/biweekly quizzes and midterm/final exams are partly related to the different intensities of study-group interactions during the regular class period (where tutorials are held) and the exams period (where lectures and tutorials are cancelled).

Among non-Beca18 students, we also find evidence of heterogeneous effects in household income: those with relatively low incomes obtain the highest academic gains if assigned to a mixed group. In terms of group composition, we find that mixed study-groups seem to influence short-term academic outcomes if they contain two Beca18 fellows but not if they contain only one. This suggests that there may be an optimal degree of diversity in small study-groups.

In the medium term (second academic term) we find that relatively low ability non-Beca18 students still benefit from having been assigned into a mixed group during their freshman semester.

[5] So even if Beca18 are of lower ability than Non-Beca18 students, the later can still benefit from mixed-group assignment.

However, we do not find statistically significant effects for Beca18 fellows in this case (although the coefficients size and sign are similar to those observed in the freshman term).

We also exploit qualitative as well as indirect, ex-post, quantitative evidence related to our study context to identify the mechanisms related to our short-term results. This evidence points to two potential underlaying mechanisms: i) differences in peers' academic quality and ii) differences in peers' attitude towards effort and cooperation, Accordingly, Beca18 fellows in mixed study-groups likely benefit from the relatively higher academic ability of their non-Beca18 peers; while non-Beca18 students in mixed study-groups plausibly benefit from the relatively higher effort and cooperation displayed by Beca18 fellows.

There is scarce evidence on educational peer effects in higher academic institutions in developing countries. There is also limited evidence on educational peer effects related to the academic interactions among college students from opposite socioeconomic status; particularly in the context of recent government interventions promoting social inclusion in higher education. Most of the college peer effects studies in the literature focus on how the academic characteristics of your peers affect your own performance; and there are also some studies on academic peer-effects related to race and gender. However, most studies analyze post-secondary academic contexts in developed countries and pay limited attention to the role played by socioeconomic status at influencing academic peer effects (for an extensive review of this literature see Sacerdote (2011)).

As in our paper, several others have exploited exogenous variation in group assignment to estimate peer effects at college. Among these, we have the seminal papers by Sacerdote (2001), Zimmerman (2003), Foster (2006), Stinebrickner and Stinebrickner (2006) and Lyle (2007). While exogenous group assignment helps researchers to deal with the endogeneity and self-selection issues related to group conformation; these studies mostly exploit students' assignment to roommates, dorms, and residences. However, as pointed by Carrell et al. (2009), roommates are just a small subset of an individual's peer group and may not constitute a relevant "academic interactions" or "reference" group; which can result in imprecise estimates. In our paper we capture a relevant "reference" group, as UDEP freshman students within the same study-group are expected to intensively interact in several of their first term courses. In this sense, our study closely relates to Carrel et al. (2009), who exploit random assignment of cadets to squadrons in the United States Air Force Academy, in which they closely interact during their freshman year.

Stinebrickner and Stinebrickner (2006) also show that low-income college students benefit from being paired with a high-income individual. They suggest that this finding may be associated with low-income students being of inferior study skills or having weaker beliefs about the importance of education; but do not provide qualitative or quantitative evidence on this matter for their estimation

sample. In our paper we show that not only low-income, but also certain high-income students benefit from being paired with an individual from the opposite socioeconomic status; and provide qualitative and indirect quantitative evidence on the potential mechanisms behind our findings.

In terms of the potential mechanisms behind student academic interactions, particularly those related to peers' attitudes towards effort and cooperation; our paper somehow connects to the recent work by Golsteyn et al. (2017). They exploit random assignment to small class sections at Maastricht University School of Business and find that students tend to perform better in the presence of more persistent peers (as well as in the presence of more risk-averse ones).

In the context of low income countries, Rao (2013) also studies the peer effects resulting from the school interactions between wealthy and poor students. He exploits a policy change in India forcing private schools to reserve a proportion of their seats for socioeconomically disadvantaged children. Rao's exploits the policy generated variation in class composition within cohorts and schools as well as an instrumental variables (IV) strategy related to whether students are usually assigned to study-groups in alphabetic order. Although Rao's paper mainly focuses on social preferences and discriminatory behavior; it also explores impacts on learning outcomes, finding marginally significant negative effects on wealthy students' English test scores. A relevant difference between our studies is that while Rao focuses on elementary school children, we focus on college individuals. Also, while he relies on an IV strategy to identify the peer effects related to working closely with a poor student, we have precise knowledge of study-group composition. Furthermore, we have detailed information related to students' pre-college academic performance and family income and are able to test for heterogeneous effects. Finally, while Rao's paper focuses only on wealthy pupils', we also study the academic outcomes of poor students.

In some sense, our paper complements the findings in studies like Carrel et al. (2013), as we also provide evidence on whether the active sorting of individuals into study-groups can be used by college and social programs officials to improve the performance of disadvantaged students.

The rest of the paper develops as follows: Section 2 describes the Beca18 program. Section 3 discusses the institutional context and the RCT implementation. Section 4 presents the empirical approach and the short-term results. Section 5 discusses potential underlying mechanisms. Section 6 discusses empirical evidence on the medium-term academic impacts. Finally, Section 7 concludes.

2. The Beca18 Program

Beca18 is a social inclusion program in higher education implemented by the Peruvian Ministry of Education (MINEDU by its Spanish initials) since 2012. It provides full scholarships to high school graduates who qualify as poor according to SISFOH, obtained a relatively high academic

achievement during their upper high school years (measured in terms of high school GPA[6,7]), and have been admitted into an eligible university. Up to 2015 the program assigned approximately 6,300 grants to students admitted into the top eight private universities in the country.

The average minimum monthly tuition in the top eight private colleges in Peru is US$ 400 (PEN S/. 1,400); while the average maximum monthly tuition is US$ 900 (PEN S/. 3,050). Given that the monetary poverty line for a five-member urban family is approximately US$ 480 per month (PEN S/. 1,660), the poor can hardly afford an elite private college education. Beca18 has therefore promoted a unique, more diverse academic context at several elite Peruvian private universities; as the relatively wealthy students who predominantly attend such institutions share now the same academic environment with individuals from the socioeconomically deprived classes.

Individuals interested in the Beca18 program must first apply and obtain admission into an eligible university. Eligible universities are free to determine their own Beca18 entry requirements and admission quota. Circumstantial evidence related to the program indicates that while in some universities Beca18 candidates take the same admission exam as non-Beca18 students and are subject to the same admission cutoffs; in others, Beca18 and non-Beca18 applicants are subject to different admission mechanisms and/or admission cutoffs. Anecdotal evidence also indicates that eligible universities adopted different strategies in terms of attracting Beca18 candidates. For example, while some universities implemented a unique centralized exam for Beca 18 fellows, others implemented decentralized admission tests in distant rural areas. As a result of these differences, there is a relatively high variation in terms of the take-over of Beca18 fellows relative to the total number of grants awarded among elite eligible universities; as well as in terms of the relative participation of Beca18 fellows in their total student population (this information is summarized in Table 1 for the top eight private universities in Peru).

Eligible universities provide admitted candidates an official admission letter, which they must submit to MINEDU with the rest of their application package. In the case of Beca18 applicants in the 2016 cohort, program officials verified the socioeconomic condition of admitted candidates (they must be poor according to the SISFOH) and ranked them according to a weighted average which depends on high school performance, the SISFOH index and other prioritizing factors (such as gender and area of residence)[8,9]. Considering this ranking and the number of fellowships announced in the

[6] A high school GPA higher than 15 in a 1 to 20-point scale is required.
[7] In the case of the SerPiloPaga program in Colombia, academic achievement is measured by a national level standardized exit exam (SABER11). In this context, it is possible to compare the pre-college relative differences in academic ability among beneficiaries and non-beneficiaries.
[8] https://www.pronabec.gob.pe/inicio/institucional/documentos/2014/r_directoral/rd472_2014.pdf
[9] Starting in 2016, MINEDU also requires Beca18 candidates to write a centralized exam which evaluates the candidates' math and verbal skills. However, the students in our study sample were not subject to this requirement, as they applied to the program in 2015: https://www.pronabec.gob.pe/modResoluciones/2018/directorales/rd83_2018.pdf

call for grants, by late May MINEDU publishes the final list of fellows. To keep their scholarship, fellows need to maintain an average GPA of 10.5 on a 20-point scale during their college studies.

3. Institutional Context at Universidad de Piura

UDEP consistently ranks among the top ten universities in Peru[10]. The admission exam for Beca18 candidates at UDEP place in late January, and the results are available in late February. As poor individuals from public high schools (particularly in rural areas) were expected to have a relatively lower high school academic background than high income individuals from private schools, and giving special consideration to the inclusive nature of the program, UDEP admission officials decided to set relatively lower admission cutoffs for Beca18 candidates[11]. UDEP also implemented a mandatory six months remedial program for admitted Beca18 fellows[12], which takes place from July to December (right after MINEDU announces the program beneficiaries). Upon approval of the remedial course, Beca18 fellows are promoted into the first academic term, starting in early March[13].

In the case of non-Beca18 students, they apply to UDEP during the months of August to December, when they are in the last year of high school (the school year in Peru starts in early March and ends in early December)[14]. Almost all admitted non-Beca18 students in the academic programs included in our study take a four-week summer program during late January and early February, which main objective is to evaluate their quantitative high school background[15]. Non-Beca18 students do not interact in any way with Beca18 fellows before the freshman term[16] .

Non-Beca18 students are mostly middle/upper middle-class individuals. Approximately 90% attended a private high school and less than 3% qualify as monetarily poor. They are therefore from significantly wealthier family backgrounds than Beca18 fellows. It is important to mention that the financial support provided by UDEP to relatively low income non-Beca18 students is not a function of the number of Beca18 fellows admitted (Beca18 fellows have therefore not displaced other types of financial aid offered by UDEP). Also, the Beca18 program did not affect the number of non-Beca18 admission vacancies offered at UDEP. As Beca18 is a public funded program, which can be

[10] https://mba.americaeconomia.com/articulos/reportajes/descubre-los-resultados-del-ranking-de-universidades-peruanas-2017

[11] Starting in 2018, Beca18 students and non-Beca18 fellows at UDEP will be subject to the same admission criteria as required by PRONABEC.

[12] Approximately 10% of admitted Beca18 fellows at UDEP fail the remedial program. Those who fail this remedial program must withdraw from UDEP and lose their grant; while those who pass are promoted into the first academic term.

[13] Almost a year passes from the time a Beca18 candidate takes her admissions exam and the time she commences her first college term at UDEP.

[14] There are several admission modalities for non-Beca18 students; such as direct admission for outstanding students from a selected list of private high school schools, admission based on a high school aptitude test, and standard admission based on a typical entrance test performance.

[15] This program is mandatory for all admitted non-Beca18 students in Engineering and Architecture. In Economics, some high quality admitted students are exempted. Those who pass this short summer program are promoted into the first academic term. Those who fail must take a remedial program exclusive for non-Beca18 students from March to June.

[16] Given the difference in the admission process for Beca18 and non-Beca18r students, Beca18 fellows are on average one year older than non-Beca18 ones when they start the first academic term.

discontinued or adjusted at any time, UDEP has not tied its regular admission decisions to the number of incoming Beca18 fellows.

Before classes begin, freshman students at UDEP are randomly assigned to class sections in which they must take all their first term courses.[17] Class size varies from approximately 60 students in Economics to 160 students in Engineering. It is also a standard practice among UDEP instructors to assign students to academically balanced small study-groups. That is, groups containing both relatively high and relatively low pre-college academic ability students[18].

To obtain a better understanding of the academic effects of students' interactions in the context of the Beca18 program, we approached UDEP authorities and proposed them to randomly assign students in the 2016 cohort to academically balanced study-groups of different socioeconomic composition (defined in terms of the individual Beca18 status). We were authorized to implement our intervention in five academic sections in the programs of Architecture, Engineering, and Economics, containing approximately 250 Beca18 and 320 non-Beca18 individuals[19]. Depending on the academic program, our study-groups were implemented in two to four courses. All these were quantitative (math) related subjects, such as Introductory Calculus, Analytical Geometry and Linear Algebra[20].

Individuals in the same study-group were expected to work together in weekly/biweekly held tutorial labs. During tutorial labs students get together in their study-groups to review the material covered in lectures and prepare in this way for their weekly/biweekly individual quizzes. Tutorial labs are however not held during midterm exams and conclude one to two weeks before final exams.

Students were not forced or encouraged in any way to use the same study-groups in other contexts, such as when preparing for exams outside tutorials, studying for other courses, or solving individual homework. Although students were encouraged to attend tutorials, attendance was not enforced. Students were also allowed to move freely around the tutorial lab classroom and ask questions or discuss problems with members of other groups. In this regard, our intervention did not force students to exclusively interact with those in their study-group; it only influenced the intensity of their academic interactions with individuals from the opposite socioeconomic status. It is then

[17] Starting in the second term, students are free to choose their course sections.
[18] Student assignment to study-groups is usually performed using a pre-college academic performance indicator; and they are generally composed by 3 to 5 individuals.
[19] We were also authorized to work with the 2017 cohort. However, while we were able to implement the random groups at the beginning of the academic term, our intervention was severely affected by the extreme floods that affected the city of Piura in March and April of 2017. As a result of this situation, academic activities were suspended just after one week after classes began. The suspension of academic activities lasted for one month approximately. When classes resumed in late April, several instructors decided to cancel tutorial labs (where study-groups meet) as well as to readjust the individual evaluation mechanisms (for example midterm exams were cancelled and students were given extraordinary exams by the end of the term to make-up for missing evaluations). As the academic context corresponding to the 2017 academic term extremely differed from that observed in the 2016 academic term, we decided not to include the 2017 cohort in our study.
[20] If in any of the courses in which our random study-groups were not implemented an instructor decided to form study-groups for a specific academic task, he/she was free to follow his/her own criteria to form these groups.

possible that such interactions took place outside randomly assigned study-groups. The evidence however suggests that such cases were rare. By the end of the first term we interviewed approximately 85% of students in the control group and asked them to list the four individuals which whom they interacted the most to prepare for exams and/or solve assignments. Only 7% listed an individual from the opposite socioeconomic group (defined in terms of the Beca18 or non-Beca18 status).

3.1. Random Assignment of Students to Study-Groups

Students within each class section were randomly assigned to three types of small study-groups: (i) groups containing only Beca18 fellows (unmixed Beca18 groups); (ii) groups containing only non-Beca18 students (unmixed non-Beca18 groups); and (iii) groups containing both Beca18 and non-Beca18 students (mixed groups). Groups were initially composed of four students, and randomization was conditional on keeping the groups balanced in terms of relative academic ability using a pre-college academic performance indicator (which we discuss later in detail).

Beca18 fellows and non-Beca18 students within each section were first identified, and individuals in each category were randomly separated into two large groups. Those in the first one were later assigned to mixed groups (treatment group), while those in the second were later assigned to unmixed ones (control group). In the case of academic programs of Engineering and Architecture, students in each of these large groups were again divided into two sub-groups using a pre-college academic performance indicator: a first sub-group containing those above the indicator median (which we call "relatively high academic ability" students); and a second sub-group group containing those below (which we call "relatively low academic ability" students). Assignment to study-groups in these programs then proceeded as follows. In the case of unmixed non-Beca18 groups, each group was randomly assigned two relatively high and two relatively low academic ability individuals; and a similar procedure was followed to assign students to unmixed Beca18 groups. In the case of mixed groups, while most contained two Beca18 students, a few ones only had one Beca18 student. In mixed groups with two Beca18 students, each group was randomly assigned a relatively high and a relatively low academic ability non-Beca18 student; as well as a relatively high and a relatively low academic ability Beca18 fellow. In the case of mixed groups with only one Beca18 student, each group was first randomly assigned a relatively high and a relatively low academic ability non-Beca18 individual. Then each group was assigned an extra individual from the remaining pool of non-Beca18 students. If this last student was of relatively high (low) academic ability, then a relatively low (high) academic ability Beca18 student was allocated.

We now discuss the pre-college academic performance indicator used to classify students as of relatively high or relatively low academic ability in the Engineering and Architecture programs. In the case of Beca18 fellows, the GPA obtained in a mandatory six-month remedial program (only

attended by Beca18 fellows at UDEP), is used as an indicator of their pre-college academic ability. In the case of non-Beca18 freshman students in Engineering and Architecture, all of them must attend a mandatory four-week preparatory summer course before college starts, therefore the GPA obtained in this course is used as an indicator of their pre-college academic ability[21].

In the case of Economics students, the procedure to assign Beca18 fellows to unmixed and mixed groups was the same as in the Engineering and Architecture programs; and the GPA obtained in their remedial program was also used as an indicator of their pre-college academic ability. In the case of non-Beca18 students in Economics, only those whose admission test score falls below a given cutoff are required to take the pre-college summer course. They represent 25% of the economics non-Beca18 students in the 2016 cohort. Therefore, random assignment in this case was conditional on not having more than one of these students assigned to a given group.

Later in the first term, and after we assigned students to their study-groups, we were informed that an instructor in the Economics program administered an entry math exam during the first week of classes (before tutorial s started), aimed at evaluating the high school math aptitude of new students. While this test score was not used to assign students to study-groups, we use it as a proxy for the pre-college academic ability of non-Beca18 students in the Economics program[22].

Table 2 presents detailed information for a set of available pre-treatment characteristics for non-Beca18 (relatively wealthier) students in the two types of study-groups (unmixed/control and mixed/treatment). Note that treatment and control groups are relatively similar, and there are not statistically significant differences in mean characteristics[23]. Table 3 performs the same exercise but focuses on Beca18 fellows. As we can see, individuals in the treatment and control group are also relatively similar. There is only one variable for which the difference in means is statistically significant at the 10% level (and the observed difference is relatively small in absolute size)[24].

[21] Therefore, the pre-college academic indicators of Beca18 fellows and non-Beca18 students are non-comparable.

[22] Random assignment took place before the term started; however during the first of classes and before tutorial labs started, there were some modifications class sections due to the following reasons: i) some students withdrew from UDEP, ii) some switched academic programs and iii) university administrators made last-minute adjustments to class lists due to space restrictions and/or exceptionally allowed some students to change sections. In this context, some groups had to be readjusted. To keep the original treatment status, we followed the simple rule that students initially assigned to a mixed/unmixed group will be reallocated also into a mixed/unmixed one. Due to this situation, a few groups finally contain five or three individuals.

[23] The socioeconomic data related to income and household characteristics was provided by UDEP administrators in standardized form. Unfortunately, the information for 15 individuals was missing in the files (less than 5% of the individuals in the data). It is important to mention that there is no systematic relationship between having your data missing in the files and the treatment, and data balance is not affected if we remove those individuals with missing data.

[24] We present the balance tests separately for Beca18 and non-Beca18 fellows, as for some variables such as income and family size we only have information for non-Beca18 fellows. Also, the pre-college academic indicators of Beca18 and non-Beca18 fellows are non-comparable, as they correspond to the GPA obtained in different pre-college programs. However, if we put the groups together and focus on those variables for which we have information for both groups, the treatment and control groups remain similar in terms of their observables (these results can be provided under request).

4. Empirical Results

To estimate the treatment effects of interest, we use the following linear model:

$$y_{igs} = \propto + \rho T_{gs} + x_{igs}\beta + \gamma_s + \mu_{igs} \quad (1)$$

Where y_{igs} is the academic outcome of individual "i" in group "g" in academic section "s"; T_{gs} is the treatment variable and indicates whether group "g" in academic section "s" is mixed or not; x_{igs} is a vector of individual characteristic, including gender, the student's pre-college academic indicator and academic program; γ_s is an academic class section fixed effect and μ_{igs} is an error term assumed to be correlated among individuals within the same study-group.

4.1. Study-Group Diversity and College Academic Performance

We analyze students' academic performance on two types of individual evaluations: i) weekly/biweekly quizzes and ii) midterm and final exams. Weekly/biweekly quizzes take place during the regular class period and assess the material covered in the immediately previous weeks (which is generally discussed during tutorial labs). These quizzes account for 30 to 40% of the final grade. Midterm and final exams on the other hand are cumulative evaluations which combined represent 40 to 60% of the final grade. During midterm and final exams (lasting approximately ten to fourteen days each) classes and tutorial labs are cancelled, and students must only come to campus to write their examinations. It is important to point out that tutorial labs usually end one to two weeks before final exams; and students are not given any homework, project or quiz a week before final exams. While this paper focus on the effects of study-group composition on non-Beca18 and Beca18 students separately, we start our analysis by considering all students in our sample at the same time.

In columns I to IV in Table 4 we estimate the effect of mixed study-group assignment on the (standardized) average grade in weekly/biweekly quizzes. To calculate this, first the standardized average grade in weekly/biweekly quizzes in each course in which random study-groups were implemented was obtained, and then the average across all these courses was estimated. Column I in Table 4 estimates a simple linear regression. As we can observe, being assigned into a mixed study-group improves performance on individual quizzes by 0.14 standard deviations; and the effect is statistically significant at the 5% level. In column II we control for class sections dummies and for a vector of individual characteristics (which include a Beca18 binary identifier as well as the individual's gender, academic program, and pre-college academic indicator). The estimated effect is similar to the one obtained in column I, and statistically significant at the 5% level.

In column III we estimate the same regression as in column II, but only include in the treated sample mixed groups that contain two Beca18 fellows (mixed groups including only one Beca18 fellow are excluded). The treatment coefficient is in this case slightly higher in absolute value than in columns I and II, as well as statistically significant. In column IV we only include in the treated

sample mixed groups containing one Beca18 fellow. Note that the estimated coefficient is very close to zero and lacks statistically significance. The coefficient estimated in column IV is also statistically different than the one estimated in column III at the 10% significance level.

In columns V to VIII in Table 4 we perform a similar analysis as in columns I to IV but focus on the standardized average grade obtained in midterm and final exams. To calculate this, we first obtain the standardized average corresponding to the midterm and final exams grades in each course, and then average across all courses in which random study-groups where implemented. Note that the estimated treatment effect coefficient is relatively small in all cases, and not statistically significant.

Why does mixed study-group assignment seem to affect academic performance on weekly/biweekly quizzes but not on midterm and final exams? As just mentioned, during the regular class period students frequently work in their study-groups in tutorial labs, where they review the material covered in the previous days and prepare for weekly/biweekly quizzes. On the other hand, during the exams periods (lasting 10 to 14 days each) lectures and labs are cancelled, study-groups are not required to meet, and students only show up to write their exams[25]. This context suggest that the observed results may be partly related to differences in the intensity of the academic interactions within study-groups during the regular class period and the exams period. However, this interpretation is still preliminary and, as we will show later in the paper, other mechanisms play also a critical role.

The results in columns III and IV in Table 4 also suggest that mixed study-group assignment is likely to affect academic performance if the study-group contains two Beca18 fellows but not if it only contains one. A plausible explanation for this result is that the academic interactions among Beca18 and non-Beca18 students may be more intense in socioeconomically balanced (more diverse) study-groups. This interpretation is supported by UDEP instructors, which mentioned that Beca18 fellows are relatively shy and more likely to engage in study group discussions if another Beca18 student is assigned to the group. Also, non-Beca18 students may put more emphasis on their interactions with other non-Beca18 individuals if assigned to a mixed study-group with only one Beca18 fellow than if assigned to a mixed study-group with two fellows. In such context, the nature of the academic interactions of non-Beca18 students in mixed study-groups with only one fellow will be relatively similar to that observed in unmixed non-Beca18 study-groups. The observed results may also be driven by the fact that a given Beca18 student is more likely to perceive himself as different from his peers if he is the only fellow in the study-group, which may affect his confidence and as result his academic performance. The two interpretations are however not exclusive, and we will further discuss them when we look at Beca18 and non-Beca18 students separately.

[25] Moreover, students were not forced or encouraged to work together in their study-groups during the exams periods.

4.2. Academic Treatment effects on Non-Beca18 Students

This section exclusively focuses on non-Beca18 students' academic outcomes. In Columns I to IV in Table 5 we estimate the effect of mixed study-group assignment on weekly/biweekly quizzes performance. Column I estimates a simple linear regression. The results show that non-Beca18 students in mixed groups improve their quizzes' performance by 0.17 standard deviations on average. In Column II we add individual controls and class section dummies to the estimated regression. The estimated effect is similar to the one obtained in column I, and statistically significant at the 10% level. Column III estimates the same regression as column II but only includes mixed study-groups containing two Beca18 fellows in the treated sample. The estimated effect in this case is close to that in columns I and II, and statistically significant at the 5% significance level. In column IV we only consider mixed study-groups with one Beca18 fellow. Note that in this case the estimated coefficient is close to zero and not statistically significant. We cannot however reject the null of no difference between the coefficients obtained in columns III and IV (likely because of the reduced sample size). The regressions in columns V to VIII in Table 5 focus on non-Beca18 students' midterm and final exams performance[26]. In all cases the estimated treatment coefficients are relatively small and not statistically significant.

The findings in Table 5 are similar to those in Table 4. They indicate that the short-term benefits of mixed study-group assignment on non-Beca18 students are mainly observed during evaluation periods in which the intensity of the academic interactions within study-groups is expected to be higher. They also suggest that, compared to unmixed groups, mixed ones seem to positively influence the academic outcomes of non-Beca18 students if they are assigned to a group that contains two Beca18 fellows (although in this case, we lack statistical significance to reject the null of no difference between the coefficients in columns III and IV). Interestingly, the coefficient for the treatment dummy in column IV is very close to zero (although imprecisely estimated). This observation is coherent with our previous discussion in section 4.1, suggesting that the nature of the academic interactions of non-Beca18 fellows in mixed-groups containing only one Beca18-fellow is likely to be relatively closer to that observed in unmixed non-Beca18 groups.

4.3. Heterogeneous treatment effects among non-Beca18 students

4.3.1. Heterogenous effects related to pre-college academic ability

We now test for heterogeneous effects related to non-Beca18 students' pre-college academic ability, which is measured using the (standardized) pre-college academic performance indicator discussed in

[26] We have missing data for 3 students, which represent less than 1% of the non-Beca18 students in the sample data.

detail in Section 3[27]. In all regressions in Table 6 we introduce an interaction term between this indicator and the treatment dummy. All regressions include class section dummies and control for the same individual variables included in column II in Table 5.

In columns I to III in Table 6 the dependent variable is the average performance in weekly/biweekly quizzes. In Column I the estimated interaction term coefficient is negative and statistically significant, suggesting that relatively low academic ability non-Beca18 students (those whose pre-college academic performance indicator is relatively low) benefit the most from mixed study-group assignment. In column II we only consider mixed groups containing two Beca18 fellows. As we can observe, the results are very similar to those in column I. The findings in columns I and II also show that not all non-Beca18 students benefit. The results in column I for example imply that non-Beca18 students whose pre-college academic indicator is above 1.05 standard deviations are negatively affected on their weekly/biweekly quizzes performance. These are students in the top 15% of the non-Beca18 pre-college academic performance distribution, and the average effect experienced by them is about -0.06 standard deviations. In column III we just consider mixed groups with only one Beca18 fellow. In this case neither the level nor the interaction term estimated coefficients are statistically significant.

In columns IV to VI in Table 6 we look for heterogeneous effects related to relative pre-college academic ability on midterm and final exams performance. Interestingly, the estimated interaction term coefficient is negative in all cases; and statistically significant (at the 10% level) in columns IV and V. On the other hand, the estimated level effect is close to zero and not statistically significant. These results indicate in first place that relatively low academic ability non-Beca18 students also benefit in terms of their midterms and finals performance; even though study-groups are not required to meet during the exam periods. Secondly, as in the case of weekly/biweekly quizzes, the results indicate that some relatively high ability non-Beca18 are negatively affected. It is however important to highlight that while in the case of weekly/biweekly quizzes only those in the top 15% of the non-Beca18 pre-college performance distribution are adversely affected, in the case of midterm/final exams those in the top half suffer academically from mixed study-group assignment.

In Section 5 we exploit qualitative and indirect quantitative evidence related to our study context to identify the mechanisms behind these results. As we will see in detail in Section 5, in first place this evidence suggests that on average Beca18 students have a stronger high school background than Beca18 fellows. Accordingly, non-Beca18 students in mixed study-groups are likely being assigned to study-groups of lower pre-college academic ability; which explains the observed negative

[27] For non-Beca18 engineering students, this indicator is the GPA obtained in a mandatory 4 weeks pre-college program. For non-Beca18 economics students, it is the grade obtained in an entry math test administered the first week of classes.

treatment effects among the academically top non-beca18 students. The evidence also suggests that due to their interactions with Beca18 fellows, non-Beca18 students increase their academic effort; which may explain the improved performance of relatively low academic ability non-Beca18 students in mixed study-groups.

The results in Table 6 also suggest that some of the mechanisms at play are influenced by the intensity of the within study-group interactions along the academic term. For example, the benefits of mixed study-groups reach a lower proportion of non-Beca18 students during the midterm and final exams periods, where students are not required to work together in their study-groups.

4.3.2. Heterogenous effects related to household income

In Table 7 we test for heterogeneous effects related to socioeconomic status among non-Beca18 students. In all regressions we add an interaction term between the treatment indicator and the non-Beca18 student's (standardized) household income. We also include class section dummies and control for students' individual characteristics. As we can observe in column II, the interaction term coefficient is negative and statistically significant at the 5% level; suggesting that the treatment benefits are stronger among non-Beca18 students with relatively low family incomes. Note that the estimated coefficient in column II is not affected after we control for the heterogeneous effect related to pre-college academic ability in column III. This rules out the possibility that the heterogeneous income effect is just proxying for the heterogeneous effect related to pre-college academic ability. Columns VI to VIII focus on midterms and finals. In this case, the estimated interaction term coefficients are also negative, and statistically significant in columns VI and VII.

Why are the mixed study-group benefits stronger among relatively low income non-Beca18 students? Probably, the intensity of their academic interactions with their Beca18 study-group peers within and outside tutorials is more likely to be higher than the one experienced by high income non-Beca18 students. As they are relatively closer to Beca18 fellows in the socioeconomic dimension than high income non-Beca18 individuals. However, this result can also be explained by a simple "attendance" effect. As tutorial labs attendance was encouraged but not enforced, relatively high income non-Beca18 students may be less likely to attend tutorials if assigned to a mixed study-group than if assigned to an unmixed one. This reduced attendance may negatively affect their academic performance. Unfortunately, we lack attendance data to shed more lights on this issue.

4.4. Academic treatment effects on Beca18 fellows

This section focuses on Beca18 fellows' academic outcomes. Columns I to IV in Table 8 estimate the effect of mixed study-group assignment on weekly/biweekly quizzes performance. In columns I to III the estimated effect is positive and relatively close in size to that observed in Table 4, but it is statistically significant only in column III, where just mixed study-groups containing two Beca18

fellows are included in the treated sample. In column IV we only consider mixed study-groups containing one Beca18 fellow. The estimated coefficient is in this case negative (-0.06), but not statistically significant. However, we cannot reject the null hypothesis of no difference between the coefficients estimated in columns III and IV (probably due to our small sample size).

While the treatment coefficient estimated in column IV is not statistically significant, the observed negative sign is coherent with one of the interpretations provided in Section 4.1 regarding the different treatment outcomes observed in mixed groups with two and one Beca18 fellows. This interpretation pointed out to the fact that a Beca18 student in a mixed group may be more likely to perceive himself as different from his peers if he is the only fellow assigned to that group, which may negatively affect his confidence, group engagement and as a result his academic performance. It can also be the case that isolated Beca18 individuals have less academic support from their non-Beca18 peers, as the later prefer to interact with other non-Beca18 students within the study-group.

Columns V to VIII in Table 8 focus on the average performance on midterms and finals. The estimated coefficient is positive in all cases, but not statistically significant. As discussed in detail in the previous sections, these results may be partly related to the fact that the intensity of the academic interactions within study-groups significantly decreases during the midterm and final exams periods.

4.5. Heterogeneous treatment effects among Beca18 fellows

In Table 9 we test for heterogeneous effects related to pre-college academic ability for Beca18 fellows. In all regressions we add an interaction term between the treatment dummy and the Beca18 standardized pre-college academic performance indicator (the GPA obtained in their mandatory remedial course at UDEP). In columns I to III the dependent variable is the average performance in weekly/biweekly quizzes; while in columns IV to VI it is the average performance in midterm/final exams combined.

In the first three columns in Table 9 the estimated interaction term coefficient is positive, but it is only statistically significant in to column II[28]; in which we restrict the treated sample to mixed study-groups containing two Beca18 fellows. The results in column II suggest that in terms of weekly/biweekly quizzes performance, the ones that benefit the most from mixed study-group assignment are the relatively high ability Beca18 fellows. That is, those with a relatively higher pre-college performance indicator (interestingly, this is the opposite of what we observed for non-Beca18 students). The results in column II also suggest that some fellows may be negatively affected. These are the Beca18 students in bottom 20% of the Beca18 pre-college academic performance distribution, whose pre-college academic performance indicator is lower than -0.98 standard deviations. In the next section we will discuss the potential mechanisms behind these results.

[28] In column I in Table 9 the corresponding p-value is 0.103

In the case of the regressions related to academic performance in midterm and final exams (columns IV to VI in Table 9), the estimated interaction term coefficient is also positive but relatively small in absolute size and not statistically significant.

5. Potential Driving Mechanisms

So far, our results suggest that mixed study-group assignment on average benefits Beca18 and non-Beca18 students, mainly in terms of their weekly/biweekly quizzes performance. However, they also indicate that not everyone is affected in the same manner; and that some students may be negatively affected. Among non-Beca18 fellows, those with a relatively low academic ability (as measured by a pre-college academic performance indicator) benefit the most; while those with a relatively high pre-college academic ability are likely to be adversely affected, particularly in terms of their midterms and final exam performance. Among Beca18 fellows, the opposite pattern is observed.

To obtain qualitative insights on the potential underlying mechanisms related to these specific findings, we conducted interviews with instructors, non-Beca18 students and Beca18[29] fellows at UDEP. Instructors mentioned that, in general, non-Beca18 students are academically better prepared for college than Beca18 fellows, likely because they received a superior high school education. As a result, on average non-Beca18 tend to perform better than Beca18 students during the first academic term. Instructors and non-Beca18 students also pointed out that fellows are hardworking, persistent and willing to help others[30]; qualities which may positively influence non-Beca18 students, particularly those with relatively a low pre-college academic performance.

When we pointed out to Beca18 fellows that non-Beca18 students in mixed groups improve their academic performance; they mentioned that this may be related to the fact that non-Beca18 students in mixed groups directly observe that Beca18 fellows work hard despite the many hardships they face, and as a result non-Beca18 students also work harder. Beca18 students also mentioned that fellows are in general cooperative and willing to explain to others the material covered in class. Although Beca18 students did not explicitly mention that non-Beca18 students are academically stronger than them or have a stronger high school background; they did mention that non-Beca18 students have better study environments (i.e. their homes) and stronger family support (i.e. educated parents), which positively affects non-Beca18 students' academic performance at college.

This qualitative evidence points to i) peers' academic ability and ii) peers' traits related to effort and cooperation as the potential underlying mechanisms behind our main findings. Accordingly, relatively low ability non-Beca18 students assigned to mixed-groups mostly benefit from having more hard-working and cooperative peers (their Beca18 peers); while relatively high

[29] The interviews took place approximately a year after our intervention.
[30] CIMA, one of the most important student volunteering associations in the Faculty of Engineering is led by Beca18 fellows.

ability Beca18 fellows in mixed-groups likely benefit from having higher academic ability peers (the relatively high academic ability non-Beca18 students assigned to their study-groups).

5.1. Mechanism I: Improved Peers' Academic Ability

This section explores in detail the role played by the "peers' academic ability" mechanism in our context. As mentioned in the previous sections, Beca18 and non-Beca18 individuals at UDEP are subject to different admission criteria and mechanisms[31], and their pre-college academic performance indicators are not comparable. Thus, we are unable to provide, ex-ante, direct quantitative evidence on the relative pre-college academic differences between Beca18 and non-Beca18 students[32]. Given this limitation, we rely on the comparison of the academic performance of students in unmixed (untreated) groups by the end of the first term, to provide an indirect (and imperfect) insight on the likely pre-college academic differences among Beca18 and non-Beca18 students.

While the analysis in this section is based on academic performance as observed by the end of the first term; there are two reasons why it can be nevertheless informative about the relative differences in pre-college academic ability among Beca18 and non-Beca18 students. In first place, as discussed in Section 3, Beca18 (non-Beca18) students in untreated groups rarely interact with non-Beca18 (Beca18) individuals. In this sense, the academic performance students in unmixed groups is expected to be relatively close to that which would be obtained by individuals with no academic interactions with peers from the opposite socioeconomic background. In second place, pre-college academic performance is strongly correlated with first term academic performance in college. A one standard deviation increase in the pre-college academic indicator is related to a 0.77 and 0.69 standard deviation increase in the average in weekly/biweekly quizzes for non-Beca18 and Beca18 students correspondingly. Therefore, the performance of students in unmixed groups by the end of the first academic term is expected to be strongly connected to their pre-college academic ability.

In the first column in Table 10 we use regression to compare the average performance in weekly/biweekly of Beca18 and non-Beca18 students in unmixed groups[33] during the first academic term. The results indicate that non-Beca18 students outperform Beca18 fellows by 0.16 standard deviations. In column II we only include in our estimation sample those Beca18 and non-Beca18 students who are above their group-specific pre-college academic indicator median (which we call relatively high ability ones). As we can see, the performance of the relatively high ability non-Beca18

[31] Beca18 fellows can only apply for admission to UDEP through a standard admission exam. In addition to the standard admission exam, non-Beca18 students an also obtain admission through a standardized academic aptitude test administered earlier in the year by UDEP in the main private schools in the Northern region of Peru. They also can obtain direct admission to UDEP if they belong to a school certified by UDEP and belong to their class top 30% percentile.

[32] We remind the reader that for Beca18 fellows this indicator is the GPA obtained in a mandatory remedial program exclusive for Beca18 fellows. While for non-Beca18 students, it is the GPA obtained in a 4-week summer course (for engineer and architecture students) or the grade obtained in an entry math exam (for economics students).

[33] We also control for academic program, class section dummies and gender.

students in untreated groups surpasses the performance of the relatively high ability Beca18 fellows in untreated groups by 0.26 standard deviations. In column III we perform a similar regression as in column II, but only include individuals in unmixed groups below their socioeconomic group pre-college academic indicator median (i.e. the relatively low ability ones). In this case, we did not find any statistically significant difference. This evidence suggests that, on average, relatively high ability non-Beca18 students are academically stronger than Beca18 fellows[34].

Keeping in mind the results in columns I to III in Table 10, it is important to highlight that by design mixed groups generally contain a relatively high academic ability non-Beca18 student. The results in columns I to III therefore suggest that, on average, Beca18 fellows in mixed groups have academically stronger peers than in unmixed ones; which may positively impact their academic performance. Particularly in the case of relatively high ability Beca18 fellows, who may be better prepared to grasp the benefits of closely interacting with higher ability individuals from the opposite socioeconomic status. Conversely, non-Beca18 students in mixed groups have academically weaker peers in mixed groups that in unmixed ones, which explain why some relatively high ability nonBeca18 students are negatively affected by the treatment.

To shed more light on the fact that improved peers' academic ability is the mechanism behind the improved performance of relatively high ability Beca18 fellows, in columns IV to VII we estimate separate regressions for relatively high ability Beca18 fellows, relatively low ability Beca18 fellows, relatively high ability non-Beca18 fellows and relatively low ability non-Beca18 fellows. As mentioned before, a Beca18 student is considered to be of relatively high (low) academic ability if her pre-college academic performance indicator is above (at or below) the Beca18 indicator median. Relatively high and low academic ability non-Beca18 fellows are defined in an analogous manner.

In all the regression in columns IV to VII we use as our treatment variable the pre-college academic indicator of the highest academic ability individual from the opposite socioeconomic status assigned to the group (instead of the binary indicator that captures group assignment). For example, for treated Beca18 fellows this will be the pre-college academic indicator of the non-Beca18 fellow with the highest academic ability assigned to the group. For untreated individuals, the treatment variable is set equal to zero. We also add an interaction term between our new treatment variable and the individual pre-college academic indicator. As we can observe, the treatment effect is only statistically significant in column IV, that is, among relatively high academic ability Beca18 fellows.

[34] A competing explanation is that instructors' grading is biased towards giving higher marks to non-Beca18 students than to Beca18 fellows. While we cannot fully discard it, we looked for evidence related to such type of grading bias among non-Beca18 students of different income levels. In first place, we compared the average grade in weekly/biweekly quizzes of non-Beca18 students in the highest and lowest income quintiles and did not find statistically significant differences. In second place, we estimated several regressions for non-Beca18 students in which we include household income as a control variable. In all regressions, the income coefficient was relatively small and not statistically significant.

In other words, only the performance of relatively high ability Beca18 fellows seems to be affected by the academic strength of the highest ability individual in the opposite socioeconomic group.

Note that in column IV the coefficient for the interaction term between the treatment variable and the individual pre-college performance indicator is negative. This indicates that among relatively high academic ability Beca18 fellows, an increase in the academic strength of the highest ability non-Beca18 student assigned to the group has a lower effect among those at the very top of the academic ability distribution. Overall, these findings suggest that improved peer's ability is the driving mechanism explaining the academic gains of relatively high ability Beca18 fellows in mixed groups.

The "peers ability mechanism" is also coherent with our finding that those at the top of the non-Beca18 students' precollege academic distribution are negatively affected if assigned to a mixed study-group. However, it does not provide a full answer to the observation that while during weekly/biweekly quizzes only non-Beca18 students in the top 15% of the pre-college academic distribution are negatively affected, during midterm and final exams those in the top 50% suffer academically. We hypothesize that this is related to the likelihood of study-group members getting together during the midterm and final exams periods, when they are not required to do so. Accordingly, we expect that group members from the opposite socioeconomic status are less likely to work together during the midterm and final exams than group members from the same status. In such situation, a non-Beca18 student in an unmixed study-group is more likely to meet with another relatively high ability non-Beca18 individual during the midterm and final exams, as two of them are generally allocated to an unmixed group. On the other hand, a high ability non-Beca18 student in a mixed group is, in first place, less likely to meet the high ability Beca18 in his study-group (as they are from different socioeconomic status), and in second place more likely to meet with the relatively low ability non-Beca18 in the study-group (as they are closer in the socioeconomic dimension). In such situation, a treated non-Beca18 student is more likely to have a lower academic performance during midterm and final exams than during weekly/biweekly quizzes. A similar reasoning also explains why we fail to find significant treatment effect for Beca18 fellows on midterm and final exams performance: they are just less likely to work together with the high ability non-Beca18 individual in their study-group during the exams period.

At the end of the first academic term, we implemented a survey in which we interviewed 85% of the students in our academic performance sample. We asked them to report the four individuals with whom they interacted the most during the first term to study for quizzes, exams and solve homework. While students in mixed groups are 16.5 percentage points more likely to report an individual from the opposite socioeconomic status (the control mean is 7%), they are also 18

percentage points less likely to report an individual from the same study-group (the control mean is 44%). This evidence is in line with our discussion in the previous paragraph,

5.2. Mechanism 2: Study Group Peers' Traits related to Effort and Cooperation

As just pointed out, qualitative and indirect quantitative evidence from UDEP suggest that non-Beca18 students are on average of relatively higher academic ability than Beca18 fellows. In this sense, non-Beca18 students in mixed groups are likely being assigned peers of relatively lower academic ability than in unmixed non-Beca18 groups. But if this is the case, why on average non-Beca18 students do benefit from mixed group assignment, particularly the relatively low academic ability ones? As noted earlier, Beca18 fellows are perceived as hardworking and collaborative, personality traits that may positively influence their non-Beca18 peers' academic performance. This section provides empirical evidence related to the role of peers' traits in our intervention context.

In the end-of-the-term survey just mentioned before[35] we also included a standard "Big Five" personality traits[36] questionnaire[37]. We focus our discussion on two of the "Big Five" dimensions assessed in the questionnaire: Conscientiousness and Agreeability. High scorers in the first tend to be perseverant, hardworking and self-disciplined; while high scorers in the second tend to be kind, dependable and cooperative. In first place, we use regression analysis in column I in Table 11 to compare the standardized scores obtained in the Conscientiousness dimension by Beca18 and non-Beca18 students in unmixed (untreated) groups by the end of the first academic. We find that Beca18 fellows score 0.6 standard deviations higher than non-Beca18 students. This observation is coherent with the qualitative evidence from UDEP suggesting that Beca18 students are relatively more hardworking and persistent than their non-Beca18 peers. We then compare non-Beca18 students in mixed and unmixed groups in column II. Interestingly, non-Beca18 students in mixed study-groups score 0.21 standard deviations higher than those in unmixed ones, and the effect is just marginally non-statistically significant (pvalue = 0.12).

The Conscientiousness index is based on the student's personal assessment of several sentences related to perseverance, effort and self-discipline using a 5-point scale[38]. Among these sentences, we have the following one: "I am a person who tends to be lazy". We generate a dummy variable that takes the value of one if the student agrees or strongly agrees with this sentence and use it as a dependent variable in a linear regression. The results, shown in column III, indicate that non-Beca18 students in mixed study-groups are 14 percentage point less likely to report that they are lazy

[35] In which we interview 85% of all students in our study-groups
[36] Our survey closely follows the Big Five Inventory 44 items questionnaire proposed by John and Srivastava (1999).
[37] Participation in the end of the term survey is unrelated to the treatment, and respondents in the treatment and control group are very similar in their observable characteristics (comparison Tables can be provided under request).
[38] Where 1 indicates 'Strongly Disagree" and 5 indicates "Strongly Agree".

than those in unmixed ones; and the effect is statistically significant at the 1% level. In columns IV and V, we estimated the same regressions as in column II and III, but only for Beca18 fellows. Note that in this case there are not statistically significant differences among treated and untreated individuals. Taken together, the evidence in columns I to V suggests in first place that non-Beca18 students in mixed groups have more hardworking peers than those they would have had if assigned to an unmixed one. In second place, our findings suggest that due to their study-group interactions with Beca18 fellows, non-Beca18 students become more hardworking, which results in higher academic performance. The attitudes towards effort of Beca18 fellows are not affected by the treatment.

As it was observed in Table 6 in Section 4.2, relatively low ability non-Beca18 students also seem to improve their performance during evaluation periods in which the study-group interactions in tutorial labs are not taking place. This particular result may be linked to the higher effort that is reported by non-Beca18 students assigned to mixed groups[39].

In column V in Table 10 we compare the scores obtained by Beca18 fellows and non-Beca18 students in unmixed (untreated) groups in the Agreeability dimension. As we can observe, there are not statistically significant differences in this case. We also did not find statistically significant differences among non-Beca18 students in mixed and unmixed groups (see column VI). However, when we compare Beca18 fellows in mixed and unmixed groups (see column VII), we find that those in mixed groups score 0.31 standard deviations higher than those in unmixed ones (the difference is statistically significant at the 1% level). We also explored the answers provided in each of the components of the Agreeability dimension[40], and found that this result is primarily driven by high scores in components related to cooperative behavior. So, why do Beca18 fellows appear to be more cooperative in mixed than unmixed groups? Maybe a higher cooperation is displayed by fellows as a strategy to integrate into their mixed study-group; or maybe they are more likely to recognize themselves as cooperative individuals after comparing their behavior with that of non-Beca18 students. In either case, it is plausible that this cooperative behavior positively affects the academic performance of their non-Beca18 peers, especially the relatively academically weak.

While the results in Tables 10 and 11 should be taken cautiously, they suggest that the academic strength of your peers as well as their attitudes toward effort and cooperation play a key role at explaining improved academic performance in the context of our intervention; and support the qualitative insights provided by instructors, Beca18 and non-Beca18 students at UDEP.

[39] However, this also can be related to the fact that they are still meeting during midterms and finals with the relatively high ability non-Beca18 student assigned to the group.
[40] Results can be provided upon request.

6. Second Academic Term Impacts

In this section we analyze academic outcomes during the second academic term at UDEP. We must highlight that our study-groups intervention was only implemented in the first term. We did not intervene in any way during the second one. No attempt was done to persuade students to keep working in their randomly assigned study-groups after the first term ended. Table 12A focuses on non-Beca18 students and Table 12B on Beca18 fellows.

In the case of non-Beca18 students, from the 321 individuals observed in the first term, 15 did not register for the second one (4.6% of the total). However, as we can observe in columns I and II in Table 12A, there is no relationship between the treatment and the decision to register in the second term. Moreover, the remaining individuals in the treatment and control groups are very similar, and there are not statistically significant differences in average characteristics among them[41]. For the second term we only have access to the students' GPA and the percentage of credits approved, and then focus our analysis on these two indicators. In column III in Table 11A we evaluate the treatment effect on the standardized GPA obtained by non-Beca18 students in the second academic term. Note that the coefficient is positive, although not statistically significant. We then test for heterogeneous effects related to relative pre-college academic ability in column IV. As we can see, the estimated interaction coefficient is negative and statistically significant, suggesting that relatively low academic ability non-Beca18 students benefit the most. The results in column IV also suggest that those in the top 30% of the non-Beca18 academic distribution are negatively affected. In column V the dependent variable is the percentage of academic credits approved in the second term. Note that in this case the treatment effect is positive as well as statistically significant. In column VI we test for heterogeneous effects in relative pre-college academic ability and find once again that those that benefit the most are the relatively low ability non-Beca18 students.

In Table 12B we perform the same analysis as in Table 12A but focus exclusively on Beca18 fellows. In the case of Beca18 fellows, 58 out 247 students (23.5%) did not register for the second academic term. Mainly because they didn´t obtain the required GPA to maintain their fellowship. Note however that the results in columns I and II indicate that the treatment is unrelated to whether they registered or not for the second term. For the remaining Beca18 fellows, the findings in columns III and V suggest that the treatment did not affect their second term GPA, or the percentage of credits approved. We also test for heterogeneous effects related to pre-college academic ability in columns IV and VI. While the estimated interaction term coefficient is positive, suggesting that the high ability individuals are the ones that benefit the most from mixed group assignment (as it was the case in the first term), it is not statistically significant.

[41] These results are not shown but can be provided under request.

It is important to assess our second-term findings in the light of the underlying mechanisms identified in Section 5. As discussed previously, for Beca18 fellows the key underlying mechanism behind the observed treatment effects was likely related to improved peer's academic ability. So why is this mechanism ineffective in the medium term? By the end on the second term we interviewed approximately 93% of all remaining Beca18 at UDEP and asked them to list the four individuals they interacted the most during the second academic term. Less than 1 in 10 fellows listed non-Beca18 student in this case, and there are not statistically significant differences among treated and untreated individuals (a similar result was observed for non-Beca18 students). A completely opposite figure was observed in the first term. In that case 30% of treated Beca18 fellows listed a non-Beca18 student when asked the same question; while only 10% of the untreated listed one (this difference is statistically significant at the 1% level). In this sense, the intervention has failed to generate sustained academic interactions between non-Beca18 and Beca18 students beyond the first term, affecting in this way the peers' academic ability channel that benefited Beca18 fellows.

But if the academic interactions between Beca18 fellows and non-Beca18 students are relatively weak during the second term, and there are not differences in this regard between treated and untreated individuals, why do relatively academically weak non-Beca18 students still benefit in the second term from their assignment into a mixed group in the first one? As it was noted in Section 5, non-Beca18 students in treated groups were significantly less likely to report that they are lazy individuals compared to those in unmixed ones. The findings in columns IV ad VI in Table 12.A therefore suggest that the academic interactions with Beca18 fellows in the first term have likely generated a permanent change in the attitudes towards academic effort among non-Beca18 students.

7. Conclusions

To the best of our knowledge, this is the one the very first studies that exploits experimental evidence related to the social inclusion program "Beca18" in a Peruvian elite private university. It is also among the first that aim to provide rigorous empirical evidence on the academic peer effects, and related mechanisms, that result from the intense study-group interactions among college students from extremely opposite socioeconomic backgrounds in a developing country setting.

Our results indicate that, on average, mixed study-group assignment benefits students in the short term; particularly during weekly/biweekly quizzes. We also find evidence of heterogeneous effects. Among non-Beca18 students, the ones that benefit the most are those with a relative low academic ability; while those at the top of the non-Beca18 academic distribution are negatively affected. Among Beca18 fellows, the opposite pattern is observed. Our evidence points to two driving mechanisms behind these results: i) peers' academic ability and ii) peers' attitudes towards effort and cooperation. While Beca18 students in mixed groups are likely being assigned peers of relatively

higher academic ability than the ones they would have had in unmixed ones; non-Beca18 students in mixed groups have more cooperative and hardworking peers. With respect to the "peers academic quality" mechanism, our results suggest that it is influenced by the intensity of the within study-group interactions at distinct stages of the academic term.

Interestingly our results also suggest that relatively low income non-Beca18 students are the ones who benefit the most from mixed study-group assignment. Maybe because they are more likely to closely interact with Beca18 fellows, as they are closer to them in the socioeconomic dimension

In the medium term (the second academic term at UDEP), our results indicate the relatively academically weak non-Beca18 students still benefit from having been assigned to a mixed group in the first term. However, we did not find statistically significant effects in the case of Beca18 fellows. These results suggest that among academically weak non-Beca18 students, their first term academic interactions with Beca18 fellows have generated a permanent change in their attitudes toward effort. In the case of Beca18 fellows, in a survey we implemented in the second term, less than 10% of control and treated individuals listed a non-Beca18 student among the four students they academically interacted the most during the second term (while 30% of the treated reported at least one in the first term). The reduced interactions of Beca18 fellows with non-Beca18 students during the second academic term may therefore explain the absence of statistically significant effects in their case.

Our paper has important implications for the implementation of the Beca18 program in Peru and other similar educational inclusion programs in developing countries. In first place, it clearly suggests that mixed study-group participation can benefit non-Beca18 and Beca18 students, and therefore can be considered as an option to improve the academic performance of socioeconomically disadvantaged and advantaged students (though through different mechanisms) during their early college years. In second place, our results highlight the importance of providing Beca18 fellows with the necessary academic background that may help them to take advantage of close academic interactions with wealthier college peers who may have a superior high-school preparation (properly designed academic remedial programs may play a crucial role in achieving this).

Finally, our results highlight the importance of promoting the interactions among non-Beca18 students and Beca18 fellows in settings other than academic study-groups. This may help to alleviate the negative effects observed among relatively high-ability non-Beca18 students, as well as to extend the benefits of these interactions to a more diverse set of academic and social settings, and beyond the first academic term.

References

Brunello, G., De Paola, M., & Scoppa, V. (2010). Peer effects in higher education: Does the field of study matter? Economic Inquiry, 48(3), 621-634.

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.

Carrell, S. E., Sacerdote, B. I., & West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. Econometrica, 81(3), 855-882.

Foster, G. (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. Journal of Public Economics, 90(8), 1455-1475.

Golsteyn, B., Non, A., & Zölitz, U. (2017). The impact of peer personality on academic achievement (No. 269). Department of Economics-University of Zurich.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. Handbook of personality: Theory and research, 2(1999), 102-138.

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point. The Review of Economics and Statistics, 89(2), 289-299.

Londono-Velez, J., Rodriguez, C., & Sánchez, F. (2017). The Intended and Unintended Impacts of a Merit-Based Financial Aid Program for the Poor: The Case of Ser Pilo Paga. Documento CEDE No. 2017-24

Melguizo, T., Sanchez, F., & Velasco, T. (2016). Credit for low-income students and access to and academic performance in higher education in Colombia: A regression discontinuity approach. World Development, 80, 61-77.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? Handbook of the Economics of Education, 3(3), 249-277.

Sacerdote, B. (2001). Peer effects with random assignment: Results from Dartmouth roommates. Quarterly Journal of Economics, 116(2).

Stinebrickner, R., & Stinebrickner, T. R. (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(8-9), 1435-1454.

Rao, G. (2013). Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools. Job Market Paper.

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. Educational Research Review, 5(2), 134-150.

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

Table 1: Beca18 Fellows in the top 8th Peruvian Private Universities (information corresponding to the academic year 2016)

National Ranking according to QS World Rankings 2015	University	Total Number of Beca18 Fellows	Participation %	Beca18 Fellows / Total Student Population
1	Pontificia Universidad Católica	489	7.7%	1.9%
2	Universidad Cayetano Heredia	905	14.3%	22.0%
4	Universidad del Pacífico	79	1.2%	1.8%
6	Universidad UPC	779	12.3%	15..0%
7	Universidad de Lima	13	0.0%	0.0%
8	Universidad de Piura	635	10%	10.3%
9	Universidad USIL	3285	51.7%	15.6%
10	Universidad ESAN	164	2.5%	NA
	Total Beca18 Fellows in the Top 8th private universities in Peru	6349	100%	

Sources: Memoria Institucional PRONABEC 2011-2016: http://www.pronabec.gob.pe/modPublicaciones/descarga/100milbecas.pdf

Table 2: Baseline characteristics for students assigned to treatment and control groups (non-Beca18 students)

	Control (assigned to unmixed group)	Treatment (assigned to Mixed group)	Difference (Absolute value)	p-value
Sex (1=Male)	0.58	0.56	0.02	0.76
N	151	170		
Age	17.29	17.27	0.02	0.83
N	151	170		
Pre-college standardized academic indicator	-0.03	0.05	0.08	0.52
N	151	170		
Family lives in the same city as the Campus is located (Yes=1)	0.66	0.59	0.07	0.17
N	151	170		
Standardized Household monthly income	0.06	-0.05	0.09	0.35
N	142	164		
Attended Private high school (Yes=1)	0.89	0.90	0.01	0.87
N	147	167		
Father has post-secondary education (Yes=1)	0.82	0.81	0.01	0.91
N	136	148		
Mother has post-secondary education (Yes=1)	0.78	0.77	0.02	0.84
N	142	164		
Mother works (Yes=1)	0.64	0.57	0.07	0.18
N	146	162		
Household size	4.21	4.36	0.14	0.32
N	147	166		

Table 3: Baseline characteristics of students assigned to treatment and control groups: (Beca18 students)

	Control (assigned to only Beca18 group)	Treatment (assigned to Mixed group)	Difference (Absolute value)	p-value
Student sex (1=Male)	0.50	0.53	0.03	0.63
N	116	131		
Age	18.32	18.45	0.07	0.28
N	116	131		
Pre-college standardized academic indicator	0.01	-0.00	0.01	0.96
N	116	131		
Family lives in Piura (Yes=1)	0.03	0.08	0.05	0.07
N	116	131		
Father with post-secondary education (Yes=1)	0.35	0.33	0.02	0.81
N	109	129		
Mother with post-secondary education (Yes=1)	0.22	0.23	0.01	0.84
N	110	122		
Mother Works (Yes=1)	0.13	0.11	0.02	0.73
N	92	105		

Table 4: Treatment Effects on Academic Outcomes
Full Student Sample (Beca18 and non-Beca18)

	I	II	III	IV	V	VI	VII	VIII
	Weekly/biweekly quizzes				Exams			
	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow
Treatment (individual was assigned to a mixed group)	**0.143****	**0.110****	**0.150*****	-0.016	0.045	0.014	0.028	-0.030
	(0.068)	**(0.049)**	**(0.050)**	(0.092)	(0.062)	(0.045)	(0.052)	(0.084)
Dummies for class section	NO	YES	YES	YES	NO	YES	YES	YES
Additional individual controls included	NO	YES	YES	YES	NO	YES	YES	YES
N	568	568	495	340	565	565	493	338
R2	0.01	0.63	0.65	0.65	0.00	0.58	0.59	0.62

Standard errors clustered at the group level appear in parenthesis.
In columns II to IV and VI to VIII the following individual controls are included: the pre-college academic indicator, the student academic program and gender.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.

Table 5: Treatment Effects on Academic Performance
Non-Beca18 Students

	I	II	III	IV	V	VI	VII	VIII
	Weekly/biweekly quizzes				Exams			
	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow
Treatment (individual was assigned to a mixed group)	**0.171*** **(0.094)**	**0.115*** **(0.068)**	**0.161**** **(0.073)**	0.002 (0.101)	0.029 (0.087)	-0.018 (0.067)	0.001 (0.072)	-0.081 (0.104)
Dummies for class section	NO	YES	YES	YES	NO	YES	YES	YES
Additional individual controls included	NO	YES	YES	YES	NO	YES	YES	YES
N	321	321	267	205	318	318	265	203
R2	0.01	0.60	0.62	0.64	0.00	0.54	0.56	0.59

Standard errors clustered at the group level appear in parenthesis.
In columns II to IV and VI to VIII the following individual controls are included: the pre-college academic indicator, the student academic program and gender.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.

Table 6: Heterogeneous Treatment Effects on Academic Performance as Function of Relative Differences in Academic Ability Non-Beca18 Students

	I	II	III	IV	V	VI
	Weekly/biweekly quizzes			Exams		
	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow	All	Only Mixed groups with two Beca18 fellows	Only Mixed groups with one Beca18 fellow
Treatment (individual was assigned to a mixed group)	**0.116*** **(0.068)**	**0.161**** **(0.073)**	0.006 (0.104)	-0.016 (0.067)	0.008 (0.065)	-0.074 (0.105)
Treatment * Pre-college academic performance indicator	**-0.110*** **(0.062)**	**-0.136*** **(0.074)**	-0.050 (0.071)	**-0.107*** **(0.062)**	**-0.128*** **(0.071)**	-0.057 (0.083)
Dummies for class section	YES	YES	YES	YES	YES	YES
Additional individual controls included	YES	YES	YES	YES	YES	YES
N	321	267	205	318	265	203
R2	0.62	0.62	0.66	0.55	0.56	0.59

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to VI the following individual controls are included: the pre-college academic indicator, the student academic program and gender.

Table 7: Heterogeneous Treatment Effects on Academic Outcomes as Function of Relative Differences in Household Income Non-Beca18 Students

	I	II	III	IV	V	VI	VII	VIII
	Weekly/biweekly quizzes				Exams			
	All	All	All	Only Mixed groups with two Beca18 fellows	All	All	All	Only Mixed groups with two Beca18 fellows
Treatment (individual was assigned to a mixed group)	**0.123*** **(0.071)**	**0.124*** **(0.072)**	**0.126*** **(0.071)**	**0.166**** **(0.075)**	-0.016 (0.069)	-0.015 (0.070)	-0.013 (0.070)	0.005 (0.076)
Treatment * Income indicator		**-0.174**** **(0.076)**	**-0.172**** **(0.070)**	**-0.161**** **(0.078)**		**-0.131*** **(0.070)**	**-0.129*** **(0.077)**	-0.112 (0.081)
Treatment * Pre-college academic performance indicator			**-0.114*** **(0.067)**	**-0.150*** **(0.078)**			-0.099 (0.066)	**-0.142*** **(0.074)**
Dummies for class section	YES	YES	YES	YES	YES	YES	YES	YES
Additional individual controls included	Yes	YES	YES	YES	YES	YES	YES	YES
N	306	306	306	254	304	304	304	253
R2	0.59	0.61	0.61	0.62	0.53	0.54	0.56	0.56

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to VIII the following individual controls are included: the pre-college academic indicator, the student academic program, standardized household income and gender.

Table 8: Treatment Effects on Academic Outcomes
Beca18 Fellows

	I	II	III	IV	V	VI	VII	VIII
	Weekly/biweekly quizzes				Exams			
	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed group with one Beca18 fellow	All	All	Only Mixed groups with two Beca18 fellows	Only Mixed group with one Beca18 fellow
Treatment (individual was assigned to a mixed group)	0.107 (0.099)	0.097 (0.067)	**0.123*** **(0.067)**	-0.055 (0.168)	0.061 (0.094)	0.054 (0.076)	0.049 (0.078)	0.098 (0.150)
Dummies for class section	NO	YES	YES	YES	NO	YES	YES	YES
Additional individual controls included	NO	YES	YES	YES	NO	YES	YES	YES
N	247	247	228	135	247	247	228	135
R2	0.01	0.69	0.70	0.69	0.00	066	0.65	0.69

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns II to IV and VI to VIII the following individual controls are included: the pre-college academic indicator, the student academic program and gender.

Table 9: Heterogeneous Treatment Effects on Academic Outcomes as Function of Relative Differences in Academic Ability Beca18 Fellows

	I	II	III	IV	V	VI
	Weekly/biweekly quizzes			Exams		
	All	Only Mixed groups with two Beca18 fellows	Only Mixed group with one Beca18 fellow	All	Only Mixed groups with two Beca18 fellows	Only Mixed group with one Beca18 fellow
Treatment (individual was assigned to a mixed group)	0.096 (0.067)	**0.124*** **(0.067)**	-0.058 (0.179)	0.053 (0.076)	0.048 (0.077)	0.092 (0.154)
Treatment * Pre-college academic performance indicator	0.101 (0.061)	**0.127*** **(0.065)**	0.028 (0.129)	0.054 (0.065)	0.054 (0.070)	0.057 (0.121)
Dummies for class section	YES	YES	YES	YES	YES	YES
Additional individual controls included	YES	YES	YES	YES	YES	YES
N	247	228	135	247	228	135
R2	0.69	0.70	0.69	0.66	0.66	0.69

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to VI the following individual controls are included: the pre-college academic indicator, the student academic program and gender.

Table 10: Potential Driving Mechanisms: Peers Relative Academic Ability

	I	II	III	IV	V	VI	VII
	Only individuals in untreated groups	Only individuals in untreated groups (relatively high academic ability)	Only individuals in untreated groups (relatively low academic ability)	Only relatively high ability B18	Only relatively low ability B18	Only relatively high ability non-B18	Only relatively low ability non-B18
Beca18 indicator (1 for B18 fellow)	**-0.156***	**-0.256****	0.001				
	(0.08)	**(0.125)**	(0.118)				
Treatment (1 if assigned to mixed group)							
Treatment (pre-college academic performance indicator of the highest ability individual from the opposite socioeconomic status assigned to the mixed group)				**0.238*****	0.067	-0.042	-0.071
				(0.088)	(0.185)	(0.083)	(0.172)
Treatment × Pre-college academic indicator				**-0.141****	0.138	0.032	-0.222
				(0.066)	(0.238)	(0.068)	(0.179)
Class Section Dummies and individual controls included	YES	YES	YES	YES	YES	YES	YES
N	268	135	133	120	127	158	163
R2	0.05	0.11	0.06	0.57	0.38	0.51	0.34

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to III the individual controls included are gender and academic program. In columns IV to VII we also add the student pre-college academic indicator.

Table 11: Potential Driving Mechanisms
Peers Personality Traits

	Consciousness					Agreeability		
	I	II	III	IV	V	VI	VII	VIII
	Only individuals in untreated groups	Only non-B18 Students	Only non-B18 students (I am a lazy person)	Only B18 fellows	Only B18 fellows (I am a lazy person)	Only individuals in untreated groups	Only non-B18 Students	Only B18 fellows
Beca18 indicator (1 if individual is a Beca18f fellow)	**0.589*** (0.108)**					-0.054 (0.125)		
Treatment (individual was assigned to a mixed group)		0.194 (0.125)	**-0.149*** (0.055)**	0.103 (0.117)	-0.040 (0.034)		0.047 (0.119)	**0.306** (0.129)**
Class Section Dummies and individual controls included	YES	YES	YES	YES	YES	YES	YES	YES
N	224	260	260	227	227	224	260	227
R2	0.14	0.03	0.08	0.06	0.03	0.03	0.03	0.12

Standard errors clustered at the group level appear in parenthesis. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns individual controls included are gender and academic program. In columns III to V and VII to VIII we also include the student pre-college academic indicator.

Table 12A: Medium Term Effects – Non-Beca18 Students

	I	II	III	IV	V	VI
	Individual registered in the second term	Individual registered in the second term	GPA second term	GPA second term	% credit approved in the second term	% credits approved second term
Treatment (individual was assigned to a mixed group)	0.010 (0.023)	0.010 (0.023)	0.098 (0.089)	0.106 (0.091)	**0.044*** **(0.026)**	**0.047*** **(0.026)**
Treatment * academic quality indicator		-0.000 (0.021)		**-0.169*** **(0.098)**		**-0.058**** **(0.029)**
Class Section Dummies and individual controls included	YES	YES	YES	YES	YES	YES
N	321	321	306	306	306	306
R2	0.07	0.07	0.36	0.37	0.20	0.21

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to VI the individual controls included are the student academic program, gender and the pre-college academic indicator.

Table 12B: Medium Term Effects – Beca18 Fellows

	I	II	III	IV	V	VI
	Individual registered in the second term	Individual registered in the second term	GPA second term	GPA second term	% credit approved second term	% credits approved second term
Treatment (individual was assigned to a mixed group)	0.023 (0.045)	0.023 (0.046)	-0.056 (0.098)	-0.098 (0.108)	-0.003 (0.030)	-0.016 (0.038)
Treatment * academic quality indicator		0.012 (0.048)		0.148 (0.128)		0.045 (0.034)
Class Section Dummies and individual controls included	YES	YES	YES	YES	YES	YES
N	247	247	189	189	189	189
R2	0.21	0.21	0.45	0.45	0.25	0.26

Standard errors clustered at the group level appear in parenthesis.
***, ** and * indicate statistical significance at the 1%, 5% and 10% level respectively.
In columns I to VI the individual controls included are the student academic program, gender and the pre-college academic indicator.