Lima SCHOOL OF ECONOMICS

WORKING PAPER - 007

Empowering college students through community engagement: experimental evidence from Peru

Marcos Agurto
Sandra Buzinsky
Fernando Fernandez
Javier Torres

May 2020

Empowering College Students through Community Engagement: Experimental Evidence from Peru

Marcos Agurto, Sandra Buzinsky, Fernando Fernandez and Javier Torres *

April 2020

Abstract

We exploit a randomized control trial involving 131 fellows of a higher education scholarship program, who study at the same university. Half of the students were randomly assigned to a youth community engagement initiative and acted as academic ambassadors in the diffusion of an electronic wallet in their local communities. They received training on leadership, teamwork and financial literacy. Also, their role as agents of change in their communities was constantly emphasized. They later delivered training and information sessions about the new electronic wallet to members of their local communities. Treated female students show positive effects regarding attitudes of empowerment, self-efficacy, motivation, and community engagement. On average, treated female students report being more appreciated by their community members, have a stronger sense of commitment towards their community, and report higher levels of self-efficacy. They also experience improved academic performance, measured in GPA and academic credits successfully completed. We do not find the same effects for treated male students.

JEL codes: H52, I25, I28, J16

Keywords: Youth Empowerment, College Academic Performance, Community Engagement, Beca 18, Peru

*We are grateful to Ana Costa-Ramón, Marion Monnet, Hugo Ñopo and Dijana Zejcirovic for helpful comments on earlier drafts. Agurto and Torres acknowledge Partnership for Economic Policy for financial support through the PIERI Award No. 12942. IRB approval obtained from PEP Ethical Committee on 2017-08-08.Agurto: Universidad de Piura and Lima School of Economics; e-mail: marcos.agurto@udep.edu.pe. Buzinsky: Universidad de Piura; email: sandra.buzinsky@outlook.com. Fernandez: Universidad de Piura and Lima School of Economics; e-mail: fernando.fernandez@udep.edu.pe. Torres: Universidad del Pacífico; e-mail: j.torresgomez@up.edu.pe. Agurto and Torres acknowledge Partnership for Economic Policy for financial support through the PIERI Award No. 12942. IRB approval obtained from PEP Ethical Committee on 2017-08-08

1 Introduction

Positive youth development theory asserts that all youth have the capacity to change and grow as they interact with their communities (Benson et al. (2007)). In recent years, there have been growing efforts in various fields to test this theory using real-world data on student outcomes (Damon (2004), Council et al. (2002), Youniss and Yates (1997), Kirshner et al. (2002), Attanasio et al. (2009), Bandiera et al. (2020), Chakravarty et al. (2019)). The findings in these studies widely suggest a positive effect of youth development programs on both cognitive outcomes, such as academic achievement, and non-cognitive ones, such as self-efficacy and motivation. However, few studies are precisely estimated due to lack of baseline data and weak channels for causal inference.

In this paper, we ask whether a community engagement program can boost both cognitive and non-cognitive outcomes among college students from disadvantaged backgrounds. The intervention delivered weekly training sessions on leadership, team work, personal development and financial literacy to a randomly selected group of students who expressed initial interest in the program. The role of these students as agents of change for their communities was emphasized throughout the treatment period. By the end of the program, students were sent into their communities to act as ambassadors, providing training and information, for a novel electronic wallet technology designed to improve financial inclusion. [1]

Relying on the experimental design of the intervention, we find that students exposed to the community engagement program increase both their academic performance and attitudes measuring self-efficacy, social responsibility, sense of belonging and empowerment. These findings complement the results of Heckman et al. (2006) which suggest that non-cognitive skills matter for academic performance. Moreover, our study suggests that such skills can be influenced even after childhood and adolescence, particularly among female individuals.

Interestingly, these effects are statistically significant for female students only. We speculate that this gender difference may be driven by existing stereotypes surrounding gender, which are prevalent in the Peruvian context, where males are conditioned by their

[1] In a related working paper, Agurto et al. (2020) explore the effect of our community engagement intervention on the adoption of electronic wallets in the respective communities where treated students provided information and training, and we refer the reader to this paper for further details.

environment to act as communal leaders whereas females are not[2]. Henceforth, providing females with community engagement opportunities, and properly training them for such engagements, could shift their beliefs about their capacity to act as agents of change.

Our results resonate with those analyzing the effects of training programs which specifically target economically disadvantaged youth. Attanasio et al. (2009) evaluate the effect of a training program for disadvantaged youth between the ages of 18 and 25 in Colombia, in which offers to receive the training were made randomly. The authors find that being offered the program increased both employment and wages. They also find that women experience large and statistically significant gains while the effects on male outcomes are statistically insignificant. Qualitative interviews from this study suggest that women had higher motivation levels than men and were less likely to drop out and be expelled from classes. Chakravarty et al. (2019) examine the effect of a vocational training to economically disadvantaged youth on job placement. Using a fuzzy RD design, they find that the program positively affected non-farm employment and average monthly earnings. Similarly to Attanasio et al. (2009), they also find heterogeneous treatment effects by gender; female gains are nearly double that of males. Bandiera et al. (2020) evaluate a randomized experiment that provided teenage girls with vocational training as well as information on sex, reproduction and marriage in Uganda. The authors find that the treatment increased participation in income-generating activities and reduced early childbearing and marriage rates.

While, as in these papers, we study the effects of a training program targeting low-income individuals, ours is not aimed at low-skill job outcomes, but rather to prepare youth to be agents of change in their communities. Moreover, we work with a targeted group of low-income youth with high levels of academic achievement, who can potentially lead the process of socioeconomic change in their communities. We add to this literature by identifying potential mechanisms through which our intervention affects the improved performance of treated individuals. Interestingly, in line with the conclusions of other studies discussed above, our results suggest that females obtain the greatest gains.[3]

[2]Which is exemplified by the high participation of male leaders in politics in contrast to females for example. As of March 1 2020, only 26 percent of congress is female according to Inter-Parliamentary Union (2020).

[3]Aside from the economic literature, there is evidence suggesting that community engagement may positively effect academic and labour force outcomes for youth. There is a large body of literature within the psychological sciences surrounding the effect of community service on youth empowerment and attitudes. see for example Council et al. (2002), Pittman (1991), Grimm Jr et al. (2005), Borgonovi (2008), Hart

The rest of the paper is organized as follows: Section 2 describes the institutional context, the community engagement program, and the experimental design. Section 3 discusses the empirical strategy, presents the baseline results related to academic performance and identifies potential driving mechanisms. Section 4 tests the robustness of our baseline results and, finally, Section 5 concludes.

2 Institutional Context and Experimental Design

Beca 18 is a national inclusion program funded by the Peruvian government, designed to provide post-secondary education for economically disadvantaged yet academically inclined students. To qualify for Beca 18, students must be under 22, come from families classified under incomes of poverty or extreme poverty, and pass a standardized exam with a sufficiently high grade.[4] We conduct a randomized experiment which administers a community engagement program to Beca 18 fellows attending a private university in Northern Peru. The program also includes training sessions aimed at providing students with leadership skills and financial literacy.

The community engagement program was conducted over two academic semesters within one private university in Northern Peru: Universidad de Piura (UDEP). The country has two academic semesters per year: semester-I runs from March until July, while semester-II runs from August until December. At the beginning of the 2017 semester-II, we randomly assigned interested Beca 18 students into treatment status and collect pre-treatment (semester 2017-I) academic outcomes. Subsequently, during 2017 semester-II and 2018 semester-I students received treatment. 2018 semester-II falls within the post-treatment time period, when follow up questionnaires are administered to students. Aca-

et al. (2007), Reinders and Youniss (2006), Clary and Snyder (1999), Zeldin et al. (2008). In the educational sciences, two studies are worth mentioning. Kirshner et al. (2002) examine the effect of a youth community engagement program on youth access and credibility to local leaders. The authors place students in a role where they investigated the needs of youth in their neighborhoods, and used their findings to influence local policies. The authors find that treated students observe positive effects with respect to self-efficacy and feelings of community engagement based on qualitative questionnaires. In a related paper, Booth and Gerard (2011) examine the relationship between self-esteem and academic success. They find that a student's self esteem at the beginning of the year has positive effects on academic test scores. When the authors consider differential gender effects, results are insignificant for male students. However, both studies pay little attention to selection into the program, and henceforth are likely affected by selection bias.

[4]In other words, Beca 18 is a combination of a need-based scheme with a merit-based component. The word "Beca" means scholarship in Spanish.

demic results are collected for every student in each semester of the experiment.[5]

A total of 131 Beca 18 fellows expressed initial interest in the program (out of approximately 500 Beca 18 fellows at the university), 68 were randomly assigned into the treatment group and 63 into the control group. Treated students attended training sessions on a weekly basis over terms 2017-II and 2018-I.[6] Table 1a shows the average characteristics of the Beca 18 fellows within our sample before treatment. We observe that students on average have a GPA of around 13.5 (out of 20 possible points), which appears to be similar for males and females. More so, a large proportion of the sample are from within the region of Piura, capturing about 60 percent of the whole sample group. This sample is also highly academically oriented, with approximately 62 percent of the full sample enrolled in engineering which is the only STEM[7] program at the university and requires the highest math achievement[8]. Finally, we observe that these students register into 24 credits on average per semester, with little variation across the whole sample. This indicates that students in our sample endure similar work loads.

Through the duration of the program, course contents varied from session to session, covering topics such as leadership, public speaking , working in team settings, the importance of body language, as well as other soft skills. Students were also given financial literacy lectures regarding the importance of savings accounts, the benefits of micro-credits, the importance of financial inclusion and information about interest rates. From the early stages of the program, students were told that the objective of the training sessions was to prepare them to be sent into their communities to initiate socioeconomic change, by way of promoting financial inclusion. Students were also told that they were going to be provided with the specific details of their financial inclusion intervention by the final stages of the training sessions.

In the final stages of the program, students were told that they were in charge of providing information and training about the new Peruvian electronic wallet, BiM, to their household network. BiM was designed and launched by Pagos Digitales Peruanos (PDP),

[5]Therefore, we obtain four data points per student.

[6]Training sessions took place on a weekly basis from late September till early November 2017 and from early May till early July 2018

[7]Sciences, Technology, Engineering, and Mathematics

[8]For instance the university has an entrance exam called the PAE, which sorts students into programs based on their relative math and verbal performance. To enter into engineering requires the highest threshold of math achievement.

a company founded by the Peruvian Banking Association, with the main purpose of promoting financial inclusion. BiM operates from any type of cellphone, even basic ones, making it more accessible to low-income communities. At the end of semester 2018-I, students went into their communities to fulfill their role as agents of change. The program paid for the students transportation costs to their communities and provided them with a 60 Peruvian soles stipend (approximately $18 USD).[9] It is important to highlight that the community engagement program was designed and implemented by Universidad de Piura in collaboration with PDP.

By providing students with leadership skills, community-level engagement and contribution, we hypothesize that there should be an improvement in attitudes of self-efficacy, empowerment, and sense of belonging, as students are encouraged by the program to act in a position of leadership in their communities to improve social inclusion. We also believe that this change in attitudes may influence students' academic performance.[10] We use survey and administrative data to empirically test for such effects. Exit surveys were conducted by the end of semester 2018-II to assess attitudes of self-efficacy, motivation, community engagement, empowerment and sense of belonging.[11] UDEP administrative records of students' GPA are used to asses the program impacts on academic performance.[12] We were also given access to academic pre-treatment data, which includes the student's GPA, credits taken, credits passed, year of studies and academic program during semester 2017-I, as well as the student's region of origin and gender. We, however, do not have pre-treatment measures of the students' attitudes mentioned above.

Table 1b reports the pre-treatment differences in mean characteristics between the treatment and control groups, with standard errors reported in parentheses. As we can observe in the first column in Table 1, the randomization successfully distributed characteristics across the treatment and control branches overall. If we look at genders separately in columns two and three, we can note that among females the pre-treatment GPA is almost half a point higher for treated than control ones, while for males the opposite pattern is observed; however, the observed differences are not statistically significant. In column three we can also observe that the variable Freshman Indicator, which is equal

[9] Later in the paper we discuss that our results are not likely to be driven by this monetary incentive.

[10] We later provide evidence that these results only holds true for female students.

[11] The response rate for the survey data is 93.2 %

[12] The GPA data was provided for all students who registered in the corresponding academic term: 98.5% of students in the 2018-I academic term and 96.2% in the 2018-II academic term.

to one if the student is in her first semester in the pre-treatment period, is unevenly distributed across treated and non-treated females; there is a higher proportion of freshman students in the treated group, with a difference significant at the 10% level. This difference may be driving the observed differences in GPA among treated and control females. In fact, within our sample, freshman students have much higher pre-treatment GPAs than the rest of students; but after their freshman term, their GPAs revert back to the mean.[13] In row four of Table 1, we estimate the mean difference in GPA among treated and control individuals conditioning on pre-treatment freshman status. As we can observe, the estimated difference in pre-treatment GPA drops considerably, and as before, remains statistically insignificant. Taking this into account, we will also report our baseline results conditioning on freshman status.

3 Empirical Strategy and Baseline Results

3.1 Empirical Strategy

Our empirical strategy hinges on the random assignment of Beca 18 fellows to treatment and control groups, conditional on the student showing initial interest in the youth community engagement program. In this context, we can interpret the mean differences in outcomes between the treated and non-treated as the causal effect of the program for those who expressed their interest in participating in it. As the period of treatment occurred throughout semesters 2017-II until 2018-I, we expect there to be treatment effects on students' performance as early as the end of semester 2018-I. We therefore estimate the program effects on academic performance for semesters 2018-I and 2018-II. We also estimate the program effects on attitudes of empowerment and social responsibility; and for this we use the results of a survey implemented by the end of the academic semester 2018-II . We first use equation (1) below to estimate the treatment effects of interest:

$$Y_i = \alpha + \beta T_i + X_i^{'}\gamma + \epsilon_i \tag{1}$$

[13] In the Section 4 we provide a detailed context based explanation for the higher pre-treatment GPA observed for freshman students relative to more senior ones

Where Y_i is the outcome variable for student i, T_i is a dummy variable which is equal to 1 if the student belongs to the treatment group, and X_i is a vector of individual-level controls. We vary the inclusion of control variables in our empirical strategy to test the robustness of our model; this includes conditioning on a student's freshman status, as well as controlling for faculty and region of origin fixed effects. In all our regressions we estimate robust standard errors.

In the case of academic performance, we provide estimates of the treatment effects using an ANCOVA regression, which allow us to directly control for pre-treatment academic outcomes. Our choice to use an ANCOVA estimation strategy over other quasi-experimental methods, such as a Difference in Differences (DID) estimation is driven by the nature of our outcome variables, which observe a low level of auto-correlation among the freshman student sample, which would create biased estimates using a DID methodology but can be accurately estimated by ANCOVA which analyzes the degree of past and future outcomes actually observed in the data. This is discussed in greater detail in section 4. Our ANCOVA regression is represented by equation (2) below:

$$Y_i = \alpha + \beta T_i + \rho Y_{i,t-1} + X_i^{'}\gamma + \epsilon_i \tag{2}$$

Where Y_i is the outcome variable for student i and $Y_{i,t-1}$ is the outcome variable in the previous period, prior to treatment. As the only pre-treatment period in our experiment is 2017-I, we control for 2017-I academic outcomes in all ANCOVA specifications. Finally, as in equation (1), T_i is a dummy variable which is equal to 1 if the student belongs to the treatment group, and X_i is a vector of individual-level controls. We again estimate robust standard errors.

3.2 Baseline Results

Table 2 presents baseline results for the program effects on semester 2018-I academic performance. First, we focus on the treatment effects on student's Grade Point Average (GPA), which can range from 0-20 points, and is a function of the student's academic performance for all courses taken in a given semester. Secondly, we study the program effects on percentage of credits approved, which equals the number of credits corresponding to

all courses for which a passing grade was obtained, divided by the total credits a student has registered in. The treatment effects are shown for the full sample population, and then separated by gender. Column 4 reports the p-values for the difference in coefficients test between male and female students.

Panel A in Table 2 reports the results of a simple OLS specification with no controls. Interestingly, for both dependent variables the effects are only statistically significant for female students. Specifically, the program increased female GPA in semester 2018-I by 0.799 grade points relative to females in the control group, which represents a 5.8% increase with respect to the control mean GPA (13.62 points). Panel A also reports a significant treatment effect on credits approved for females. Treated female students on average approved 12.6 percentage points more credits than females in the comparison group, which represents a 14.5% increase from the control mean.

Panel B reports the treatment effects conditioning on pre-treatment GPA, that is, it estimates an ANCOVA regression. While treatment effects are still statistically significant for the female students only, the magnitude and significance of the estimates decreases in comparison to those in Panel A. The estimated treatment effect on GPA in Panel B is 0.517 points, which is significant at the 10% level and reflects a 3.8% increase in GPA with respect to the control mean. We also can observe in Panel B that the estimated increase on the percentage of credits approved for treated female students is about 9.72 percentage points, an 11.2% increase from the control mean.

In Panel C, in addition to pre-treatment GPA, we also control for Freshman Status, which addresses the lack of balance in this category between treated and non-treated female students. As it can be observed, the estimated treatment effect on females 2018-I GPA in Panel C is 0.656 points, which is relatively close in size to the estimated effect in Panel A. This effect is significant at the 10% level and constitutes a 4.8% increase from the control mean. Panel C also reports a treatment effect on credits approved for female students of 10.3 percentage points. This effect is statistically significant at the 5% level and represents an increase of 11.9% credits approved with respect to the comparison group. In the specification corresponding to Panel D, we control for pre-treatment GPA, freshman status, and also allow for faculty [14] and region of origin[15] fixed effects. Consistent with the previous specifications, Panel D shows a positive and statistically significant treatment ef-

[14] Engineering, Economics, and Architecture

[15] Amazonas, Cajamarca, La Libertad, Lambayeque, Lima, Piura, San Martin, and Tumbes

fect on both 2018-I GPA and percentage of credits approved for female students only, with similar magnitudes as those observed in Panel C.[16]

In column 4 in Table 2 we report the p-value for the difference in treatment effects among female and male students. The results corresponding to Panel A suggest that the difference in treatment effects observed among genders is statistically different from zero. In all other panels, while the magnitude of the effects is consistently higher for females, the difference in effect size is not statistically significant. Nevertheless, in all but one case and despite our small sample, the p-values in panels B, C and D are relative close to the 10% significance level. Overall, the evidence in Table 2 points to the presence of positive and significant treatment effects on female academic performance. Interestingly, it also suggests that male students are robustly unaffected by the treatment.

As mentioned earlier in Section 2, treated students didn't complete their full treatment period until the end of semester 2018-I, and henceforth it is possible that there may be differential treatment effects between semester 2018-I and 2018-II. Specifically, community-level engagement didn't occur until the final stages of 2018-I; for that reason, Table 2 may not capture the full effect of treatment on academic performance. Alternatively, if there are diminishing treatment effects over time, we may observe treatment effects of smaller magnitude and significance. Panels A, B, C and D in Table 3 estimate the same regressions as their equivalents in Table 2; but focus on 2018-II academic performance outcomes. As it can be observed, the treatment effects corresponding to 2018-II GPA are relatively similar in magnitude and statistical significance to those observed for the academic semester 2018-I; and as before, they concentrate among female students. Note however that while in Table 2 there are also statistically significant effects on the proportion of credits approved among females, that is not the case in Table 3. As shown in Table 2, control individuals on average failed more courses than treated ones during semester 2018-I. Subsequently, during semester 2018-II these students are more likely to be registered in courses they have already attended, which increases their chances of passing them, and therefore washes out the differential in credits approved among treated and untreated students for semester 2018-II. Overall, the results in Table 3 confirm that the treatment effects generated by the community engagement program on female students

[16]For robustness, we have checked treatment effects with the inclusion of more stringent fixed effects, at the region and program level which decomposes each academic faculty into its respective programs. This specification loses magnitude and significance on the treatment effects, which is likely due to over-specification relative to the sample size.

persisted over time; while in contrast, male students remain consistently unaffected.

3.3 Potential Driving Mechanisms: Empowerment, Attitudes of Social Responsibility and Sense of Belonging

The community engagement program was non-academic in nature; it was primarily designed to provide students with tangible leadership and life skills, and to engage them in the problems of their communities. Accordingly, we suspect that a possible channel by which student academic performance outcomes are affected is through increased attitudes of empowerment, sense of belonging, self-efficacy, confidence, motivation, and community engagement. These non-cognitive gains can create spillovers into other areas of a student's life, such as academic performance.

Table 4 provides results from an exit survey regarding attitudes of social responsibility and empowerment. This survey was conducted at the end of semester 2018-II. Survey questions are phrased as statements, where students must choose how much they relate to each statement on a scale of 1-7; with 7 and 1 being the most and least accurate respectively. Since some of the statements have a positive connotation and some others have a negative one, to facilitate the interpretation of our results the responses are re-scaled to indicate a positive connotation for all survey questions. As in all other tables, treatment effects are distinguished between sex, with column four reporting p-values for the difference in coefficients test between male and female treatment effects.

Through the entire program treated students were constantly reminded about their role as agents of change in their communities. By the end of the program they were assigned to the role of community ambassador for a novel electronic wallet technology (BiM). A student's exposure to these particular features of the program may have positive effects on her attitudes towards her community and its problems. Table 4 verifies that this effect does in fact exist. Out of 14 survey questions related to social responsibility and empowerment, 7 present positive and statistically significant treatment effects among female students. For instance, we find a positive and statistically significant (at the 5% level) treatment effect for female students for the second statement of the survey, which states: *"I am engaged with the problems of my community"*. Treated female students agreed with this statement on average, 0.618 points more than non-treated ones (out of a scale from 1-7).

This represents a 8.8% higher rate of agreement relative to the control mean. Similarly, we observe a positive and statistically significant treatment effect on female students regarding the ninth statement related to attitudes of social responsibility, *"I have the responsibility to help other members of my community get the same opportunities which I have gotten"*. The treatment effect in this case represents an 8.9% higher rate of agreement relative to the comparison group.

Additionally, we find positive effects on social responsibility, as the responses of treated female students to statement twelve in Table 4 reveals they are more likely to view their community's problems as their own. The same can be said for the treatment effect on community sense of belonging and empowerment among female students. Answers to statements three, eleven, and thirteen show that treated female students have a higher understanding of their community's problems and have a higher valuation of their ability to address them in an impactful way. Interestingly, although they were treated in the same capacity and given the same engagement roles, we do not find statistically significant effects for any of the statements in Table 4 among male students.[17]

Overall, in comparison to non-treated female students, females in the treatment group experience positive effects on social responsibility, community engagement, empowerment, and develop a greater sense of belonging to their community. If a student feels empowered and socially responsible to create change at her community-level, it can be expected that she will be more likely to feel empowered to improve individual-level outcomes as well. In this sense, the improved performance in academic performance observed among treated female students is likely to be driven, at least in part, by the non-cognitive gains discussed in Table 4. Moreover, the proposed mechanism and its related evidence are in line with the findings in the psychological literature linking non-cognitive development with cognitive gains among the youth. Interestingly, we only find this effect to be robustly true for female students in our sample.

[17]For 11 of the 14 statements in Table 4 the treatment coefficients are relatively smaller in size for male individuals, and in 2 cases the observed difference among genders is statistically significant

3.4 Potential Driving Mechanisms: Self-Efficacy, Confidence, and Motivation

The students in our sample come from disadvantaged backgrounds and it is possible that, given their socioeconomic contexts, they may suffer from relatively low levels of self-efficacy, confidence, or motivation.[18] By providing these students with leadership training, they may strengthen these traits through changing attitudes surrounding agency within the scope of community engagement. Potentially, this could spillover into agency in other aspects of their lives; particularly, they may gain confidence and motivation to strive for better outcomes academically.

Table 5 shows the youth community engagement program effects on student self-efficacy, confidence, and motivation levels. Out of the 13 questions regarding these attitudes, two report a statistically significant treatment effect, which seems to be primarily driven by female treated students. Specifically, two statements appear statistically different for treated and control groups: *"I can resolve the majority of my problems if I put in the necessary effort"* (Statement 6), and *"I stay calm during difficult situations, because I trust my ability to respond"* (Statement 9). The first statement shows an overall treatment effect of 0.19 relative to the control group, and it is significant at the 10% level. The effect, nevertheless, concentrates among female treated students. for whom we observe a treatment effect on statement 6 of 0.237 points, which is significant at the 10% level. Regarding statement 9, there is a large treatment effect on responses to this question for the female sub-sample. In this case, treated females on average agree with this statement by 0.724 points more that control ones (a 10.3% increase), and the effect is statistically significant at the 1% level. There is no statistically significant effect for males for statement 9 and, in fact, the coefficient is negative and close to zero.

The statistically significant treatment effects observed for females in Table 5 could be due to the tendency of women to diminish and undervalue their professional skills and achievements, which is already in place by adolescence (Pajares and Schunk (2001)). We argue that for male students, the lack of treatment effect on self-efficacy and confidence is due to the fact that men are more likely to be socialized to be confident, assertive, and self-promoting, whereas women are typically socialized to be self-critical. This suggests a

[18]Self-efficacy can be defined as a personal judgment of "how well one can execute courses of action required to deal with prospective situations" Bandura (1977).

potential for larger treatment gains for women.

Following the work of Kling et al. (2007), we create a set of summary indices for the 27 outcome variables shown in Tables 4 and 5. Specifically, we combine multiple survey statements under one umbrella index to aggregate information regarding multiple treatment effect. The goal of this is to draw general conclusions about the treatment effect on non-cognitive outcomes and to improve statistical power.[19] Table 6 provides our summary index of treatment effects on non-cognitive outcomes, with robust standard errors reported in parentheses.

To understand the broader effect of the community engagement program on the non-academic outcome variables, we define seven index variables to represent different possible non-cognitive developmental outcomes affected. These index variable are: Social Responsibility, Community Engagement, Sense of Belonging, Empowerment, Self-efficacy, Confidence, and Motivation. Within each index, 2 to 5 survey statements from Tables 4 and 5 are matched to the most appropriate index.[20] We then estimate the average effect of receiving treatment on each index variable, giving equal weight to each survey statement within the index. In this manner, we obtain a more balanced and interpretable picture of the effect of the community engagement program on non-cognitive outcomes.

Interestingly, we find that the positive and statistically significant effects concentrate among statements surrounding social responsibility and community engagement. For the social responsibility index, we observe an overall positive and significant treatment effect of 0.274 points, which represents a 4% increase relative to non-treated students and is statistically significant at the 5% level. Among female students only we observe an improvement in attitudes of social responsibility of 7.5% relative to females in the comparison group (significant at the 1% level). Among male students, however, we do not find an effect statistically different from zero. This finding is common throughout the table. There is no significant effect for male students in any of the index variables presented in Table 6.

[19] Following the estimation strategy of Kling et al., our results are interpreted as the differences between treatment and control means of an index, the same as the average of treatment and control means of each components of that index so that the index can be interpreted as the average of results for separate measures.

[20] Social Responsibility Index: Table 4 statements 3, 9, 12, 13. Community Engagement Index: Table 4 statements 2, 4, 7. Sense of Belonging Index: Table 4 statements 1, 5, 8, 11. Empowerment Index: Table 4 statements 6, 10, 14. Self-Efficacy Index: Table 5 statements 1, 5, 6, 8, 10, 26. Confidence Index: Table 5 statements 4, 7, 9, 11, 12. Motivation Index: Table 5 statements 3, 13.

Attitudes of community engagement are also strikingly affected by the treatment program. Among the full-sample, treated students experience an increase in attitudes related to community engagement of 0.425 points, or 6% increase. As we can observe, this effect seems to be mainly driven by female students, who experience a treatment effect of 0.591 points, an 8.4% increase compared to non-treated females. Both of these results are significant at the 1% level.

Although we find some indication of positive treatment effects for female students, from tables 4 and 5, on attitudes of sense of belonging, empowerment, self-efficacy, motivation, and confidence, their index variables obtain results that cannot be distinguished from 0 at the 10% level. Overall, our results suggest the community engagement program has effects on specific dimensions of non-cognitive outcomes.

3.5 Ruling out Alternative Mechanisms

As previously mentioned, by the end of the academic term 2018-I, the community engagement program reimbursed participant students for the transportation costs to their communities in order to deliver the BiM training sessions. In addition, each treated student received a monetary stipend of 60 Peruvian soles to cover other travel related expenses (close to US $18). Although this reimbursement amount was not tied in any way to their performance in the program or used as an incentive mechanism, it is still possible that students' improved academic achievement is driven, at least in some part, by the monetary payment expected by them.[21] Nevertheless, we think that this is not likely to be the case. In the first place, it is not clear why this mechanism should only affect the academic performance of females and not that of males; neither why it should drive females students improved academic performance after the intervention finished (2018-II academic term), even when no additional payments were expected. Secondly, if this was the main mechanism behind our results, it is hard to explain why this should have generated a change in the attitudes that discussed in the previous two sections, and again only for female individuals.

To alleviate concerns that a monetary transfer could be driving treatment effects, we

[21] As they are low-income students this could be even more salient, as the transportation amounts provided to treated students are likely economically relevant.

include data related to the transportation amounts received by each student to empirically check the relevance of this alternative mechanism. To test this effect, we create an additional table, Appendix Table 2, which adds the variable "amount received" as an additional control to our simple OLS regression in Panel A of Table 2.[22] When we do this, we find that the treatment coefficient retains its sign and remain very similar in magnitude, but loses its statistically significance. This is likely a result of the high correlation among the treatment variable and the amount received variable. Moreover, the coefficient for the amount received variable is very close to zero, not statistically significant, and has the opposite (negative) sign as to what we would expect if monetary incentives were indeed driving our results. As an additional check, we estimate a regression among only treated individuals to examine the correlation between the GPA obtained in a given semester and the monetary amount received. This result is presented in Appendix Table 3. Again, the estimated coefficients of monetary value on academic performance variables are very small and not statistically significant. Putting together all the evidence we have observed, it is highly unlikely that our results are purely explained by the response of students to the monetary amounts paid to them in order to cover their transportation costs and related travel expenses.

4 Falsification and Robustness Checks

Since the engagement program was carried out during the semesters 2017-II and 2018-I, treatment effects should only occur after treatment has at least commenced, in 2017-II. Table 7 provides a falsification test by estimating the treatment effect on pre-treatment academic outcomes (semester 2017-I). The results confirm that there are not statistically significant effects on our outcome variables of interest, GPA and credits approved, prior to the intervention. These results are robust to different specifications, such as conditioning on freshman status in Panel B, and the inclusion of fixed effects in Panel C. Since Table 1 in Section 3 indicates a well-balanced randomization at baseline [23], it is not surprising that there are no statistically relevant effects before the treatment period.

In addition to our falsification test, we also estimate a DID regression for the treatment

[22]Note that while there is some variability in how much treated students receive as a function of travel costs to their home community, control students receive nothing.

[23]With the exception of freshman status, which was slightly higher in the treatment group at baseline.

effects of our intervention on academic performance. The results are shown in Appendix Table 1. Our DID regressions are estimated as follows:

$$Y_{it} = \alpha + \beta D_i + \gamma Post_t + \rho D_i * Post_t + \epsilon_{it} \tag{3}$$

Where Y_{it} is the outcome variable of interest for individual i in time t, where t = [0,1]. D_i is a dummy variable equal to one if student i belongs to the treatment group, while $Post_t$ is a dummy variable equal to one if t=1, or in other words, in the post-treatment period. Finally, the interaction of these dummy variables provide us with the causal effect of interest, ρ.

In the case of treatment effects on GPA, we obtain estimates which are relatively close in magnitude to those in the ANCOVA regressions in Tables 2 and 3; however, they lack statistical significance. In the case of treatment effects on credits approved, the results are also similar to those in the ANCOVA regressions in terms of size and statistical significance. Overall, the DID estimates suggest that our results are relatively robust.

Under a randomized experiment, conditional on the even distribution of characteristics between treated and non-treated groups, a simple comparison of means is sufficient specification for causal inference. However, under certain circumstances, an OLS regression may not be optimal, even within the case of a randomized experiment. In fact, McKenzie (2012) argues that when outcome variables have low auto-correlations[24] authors should control for previous levels of the outcome variable whenever possible.[25] This can be addressed through the inclusion of multiple survey rounds and estimating a DID regression to subtract pre-treatment outcomes from the estimated effect, or an ANCOVA regression to directly control for pre-treatment outcomes in the specification. To optimize the statistical power of the test, Mackenzie provides a strong case to prefer ANCOVA over DID when auto-correlations are low. In this scenario, the identifying assumption of parallel trends in DID estimation tends to be unreliable due to the high variability of the outcome variable combined with the observation of only two time periods, which leads to

[24]Meaning that they observe a high degree of variability and a low degree of predictability from period to period

[25]The problem with low auto-correlation of the outcome variable in a randomized experiment is that the estimated effects are highly time sensitive; there could be then a large confidence interval of the true effect, depending on the time of the post-treatment survey. Therefore, it is difficult to precisely estimate the true treatment effect

a large confidence interval of potential effect magnitude. Correcting for baseline data via DID may therefore over-correct for baseline differences in means which do not have much predictive power. Alternatively, the ANCOVA produces more accurate results with lower standard errors by adjusting the degree of correction of baseline differences in means between treatment and non-treatment groups according to the degree of correlation between past and future outcomes which are actually observed in the data[26].

Although our academic performance outcome variables, GPA and Percentage of Credits Approved, are typically highly auto-correlated on the individual level, we observe a low auto-correlation in GPA among freshman students in our sample in comparison to non freshman students. The GPA auto-correlation among students who start the program immediately after their freshman semester is 0.262, while the GPA auto-correlation among the rest of students is 0.703, a much higher and predictable level. During the academic semester 2017-I, the region which the university is located was affected by extreme rain and floods, and as a result classes were suspended for two months. Beca 18 students were among those that were severely affected by this event. Special attention was then given by the university to Beca 18 students, particularly among freshman students, and a more flexible grading scheme was administered during semester 2017-I to these students, in response to the local context. This situation may offer a reasonable explanation of the low degree of correlation in GPA among freshman students in our sample.[27] As freshman students represents 23% of our full sample, and since there is no disadvantage of using ANCOVA over differences in differences, we maintain the former as our preferred specification.

5 Conclusion

We analyze the effects of a youth community engagement program targeted towards underprivileged university students in Northern Peru. We add to the existing literature regarding community engagement positive effects on youth development by providing quantitative analysis based on a random control trial experiment, and identifying casual

[26]This benefit of ANCOVA over DID is especially pronounced under the presence of multiple survey rounds, where authors are better able to capture auto-correlation in the outcome variable

[27]Another plausible explanation for the observed low level of GPA correlation among freshman students is that, in general, the first semester is relatively easier than the more advanced ones

estimates of the magnitude community engagement can impact youth cognitive and non-cognitive functions, and attitudes of empowerment.

Our results indicate that the program has significant and positive effects on academic performance outcomes for female students, yet no distinguishable effect for male students, with results that are robust to the inclusion of a vector of controls and specifications. We estimate roughly a 5% increase in GPA scores for female students who receive the community engagement program, in comparison to females who did not. Similarly, we find positive and significant treatment effects for the percentage of credits approved in a semester for female students, with an estimated average treatment effect of 11.9% increase in credits approved. We argue that the driving mechanism behind these academic improvements is through increased attitudes of empowerment, social responsibility, community engagement and sense of belonging. We also find weak evidence in self-efficacy gains among treated female students.

Female students are empowered disproportionately to male students. We find strong evidence that females experience positive effects on self-reported attitudes of empowerment within their community, while there is no discernible difference between male treated and non-treated students. This differential treatment effect between men and woman sheds light on gender stereotypes that surrounds leadership; while men are socialized to be self-promoting and assertive, women are not.

It's possible that male students do not experience significant treatment effects because they are already conditioned to believe they are capable of leadership, while females may receive larger treatment effects if practical leadership training and community engagement instills confidence and assertiveness. This paper provides promising results to policy-makers showing that female outcomes can improve with low-cost training interventions.

References

Attanasio, O., Kugler, A. D., and Meghir, C. (2009). Subsidizing vocational training for disadvantaged youth in developing countries: evidence from a randomized trial.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., Sulaiman, M., et al. (2020). Women's economic empowerment in action: Evidence from a randomized control trial in africa.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191.

Benson, P. L., Scales, P. C., Hamilton, S. F., and Sesma Jr, A. (2007). Positive youth development: Theory, research, and applications. *Handbook of child psychology*, 1.

Booth, M. Z. and Gerard, J. M. (2011). Self-esteem and academic achievement: a comparative study of adolescent students in england and the united states. *Compare: A Journal of Comparative and International Education*, 41(5):629–648.

Borgonovi, F. (2008). Doing well by doing good. the relationship between formal volunteering and self-reported health and happiness. *Social science & medicine*, 66(11):2321–2334.

Chakravarty, S., Lundberg, M., Nikolov, P., and Zenker, J. (2019). Vocational training programs and youth labor market outcomes: Evidence from nepal. *Journal of Development Economics*, 136:71–110.

Clary, E. G. and Snyder, M. (1999). The motivations to volunteer: Theoretical and practical considerations. *Current directions in psychological science*, 8(5):156–159.

Council, N. R. et al. (2002). *Community programs to promote youth development*. National Academies Press.

Damon, W. (2004). What is positive youth development? *The Annals of the American Academy of Political and Social Science*, 591(1):13–24.

Grimm Jr, R., Dietz, N., Spring, K., Arey, K., and Foster-Bey, J. (2005). Building active citizens: The role of social institutions in teen volunteering. youth helping america. *Corporation for National and Community Service*.

Hart, D., Donnelly, T. M., Youniss, J., and Atkins, R. (2007). High school community service as a predictor of adult voting and volunteering. *American Educational Research Journal*, 44(1):197–219.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, 24(3):411–482.

Kirshner, B., Fernández, M., and Strobel, K. (2002). Youth community engagement: A sociocultural study of participatory action research. *American Educational Research Association, New Orleans, LA*.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more t in experiments. *Journal of development Economics*, 99(2):210–221.

Pajares, F. and Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception*, 11:239–266.

Pittman, K. J. (1991). Promoting youth development: Strengthening the role of youth service and community organizations.

Reinders, H. and Youniss, J. (2006). School-based required community service and civic development in adolescents. *Applied Developmental Science*, 10(1):2–12.

Youniss, J. and Yates, M. (1997). *Community service and social responsibility in youth*. University of Chicago Press.

Zeldin, S., Petrokubi, J., and MacNeil, C. (2008). Youth-adult partnerships in decision making: Disseminating and implementing an innovative idea into established organizations and communities. *American Journal of Community Psychology*, 41(3-4):262–277.

6 Tables

Table 1a. Raw Mean Pre-Treatment Characteristics

	I	II	III	IV	V	VI
	Full Sample		**Males Only**		**Females Only**	
	(Treatment)	(Control)	(Treatment)	(Control)	(Treatment)	(Control)
GPA in semester 2017-i	13.50	13.62	13.49	14.21	13.51	13.03
	(0.211)	(0.266)	(0.307)	(0.370)	(0.295)	(0.357)
Cumulative GPA in semester 2017-i	13.67	13.74	13.57	14.15	13.76	13.33
	(0.195)	(0.254)	(0.288)	(0.360)	(0.268)	(0.349)
Freshman Indicator	0.206	0.159	0.0909	0.194	0.314	0.125
	(0.0494)	(0.0464)	(0.0508)	(0.0721)	(0.0796)	(0.0594)
GPA in semester 2017-i,	13.18	13.17	13.35	13.59	12.96	12.78
Controlling for Freshman Status	(0.223)	(0.251)	(0.317)	(0.322)	(0.310)	(0.372)
Cumulative GPA in semester 2017-i,	13.39	13.31	13.43	13.51	13.33	13.12
Controlling for Freshman Status	(0.206)	(0.238)	(0.297)	(0.298)	(0.283)	(0.369)
% credits approved in semester 2017-I	0.880	0.881	0.880	0.906	0.880	0.855
	(0.0234)	(0.0227)	(0.0328)	(0.0261)	(0.0339)	(0.0369)
% credits approved in semester 2017-I,	0.852	0.858	0.868	0.883	0.832	0.834
controlling for Freshman Status	(0.0282)	(0.0259)	(0.0354)	(0.0308)	(0.0459)	(0.0408)
Number of credits the individual	24.10	23.66	24.12	24.03	24.09	23.29
registered during term 2017-i	(0.209)	(0.268)	(0.298)	(0.389)	(0.297)	(0.363)
Individual is from Piura	0.618	0.603	0.636	0.581	0.600	0.625
	(0.0594)	(0.0621)	(0.0850)	(0.0901)	(0.0840)	(0.0870)
Individual registered in an engineering program	0.618	0.651	0.727	0.710	0.514	0.594
	(0.0594)	(0.0605)	(0.0787)	(0.0829)	(0.0857)	(0.0882)
Individual registered in the economics program	0.338	0.302	0.212	0.258	0.457	0.344
	(0.0578)	(0.0583)	(0.0723)	(0.0799)	(0.0854)	(0.0853)
Sex (female=1)	0.515	0.508	-	-	-	-
	(0.0611)	(0.0635)	-	-	-	-

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column four displays a p-value test of the difference in treatment effect between males and females.

Academic performance outcome variables are GPA, which takes values 0-20, and % of Credits

Approved, which is the total credits a students obtains from a given semester divided by the total credits they have applied to complete.

Table 1b. Differences in Mean Pre-Treatment Characteristics by Treatment Status (Treated - Non-Treated Sample)

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(131)	(64)	(67)	(Females - Males)
GPA 2017-I*	-0.119	-0.715	0.477	0.070*
	(0.339)	(0.480)	(0.463)	
Cumulative GPA 2017-I*	-0.073	-0.583	0.430	0.108
	(0.320)	(0.461)	(0.440)	
Freshman Indicator	0.047	-0.103	0.189*	0.026**
	(0.068)	(0.088)	(0.099)	
GPA 2017-I, Conditioning on Freshman Status	0.011	-0.239	0.188	0.513
	(0.336)	(0.452)	(0.484)	
Cumulative GPA 2017-I	0.076	-0.085	0.212	0.629
Conditioning on Freshman Status	(0.315)	(0.420)	(0.465)	
% Credits Approved 2017-I	-0.0002	-0.026	0.025	0.429
	(0.033)	(0.042)	(0.05)	
% Credits Approved 2017-I	-0.005	-0.015	-0.002	0.859
Conditioning on Freshman Status	(0.0382)	(0.047)	(0.061)	
Number of credits registered 2017-I	0.442	0.089	0.795*	0.292
	(0.340)	(0.490)	(0.469)	
From Piura Region	0.015	0.056	-0.025	0.637
	(0.086)	(0.124)	(0.121)	
Individual registered in an	-0.033	0.017	-0.079	0.559
engineering program	(0.085)	(0.114)	(0.123)	
Individual registered in the	0.037	-0.046	0.113	0.319
economics program	(0.0821)	(0.108)	(0.121)	
Sex (female=1)	0.007	-	-	-
	(0.088)	-	-	-

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports the difference in baseline characteristics between the treatment and control group, among the full sample, males only, and females only. Column four reports a p-value test between male and treatment differences at baseline

*GPA shows the average grade point average for student i among his courses in a given semester, in this case semester 2017-I.

GPA takes score from 0-20. Cumulative GPA is a combination of all semester GPA scores that a student has received.

Table 2. Treatment Effects on 2018-I Academic Performance

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(129)	(64)	(65)	(Females - Males)
Panel A. Simple OLS				
GPA 2018-I	0.209	-0.383	0.799**	0.054*
	(0.314)	(0.486)	(0.386)	
% Credits Approved	0.0477	-0.031	0.126**	0.039**
	(0.039)	(0.055)	(0.053)	
Panel B. Conditioning on Pre-Treatment GPA				
GPA 2018-I	0.250	-0.013	0.517*	0.293
	(0.261)	(0.412)	(0.308)	
% Credits Approved	0.051	-0.00007	0.097**	0.156
	(0.035)	(0.051)	(0.047)	
Panel C. Conditioning on Pre-Treatment GPA and Freshman Status				
GPA 2018-I	0.316	-0.085	0.656*	0.138
	(0.254)	(0.387)	(0.338)	
% Credits Approved	0.055	-0.006	0.103**	0.115
	(0.035)	(0.051)	(0.05)	
Panel D. Including Region and Faculty Fixed Effects				
GPA 2018-I	0.261	-0.130	0.666*	0.105
	(0.280)	(0.379)	(0.388)	
% Credits Approved	0.052	-0.009	0.141**	0.033**
	(0.039)	(0.051)	(0.058)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column four displays a p-value test of the difference in treatment effect between males and females.

Academic performance outcome variables are GPA, which takes values 0-20, and % of Credits Approved, which is the total credits a students obtains from a given semester divided by the total credits they have applied to complete

Table 3. Treatment Effects on 2018-II Academic Performance

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(126)	(62)	(64)	(Females - Males)
Panel A. Simple OLS				
GPA 2018-II	0.217	-0.311	0.744**	0.064*
	(0.291)	(0.446)	(0.367)	
% Credits Approved	0.007	0.011	0.004	0.931
	(0.035)	(0.049)	(0.053)	
Panel B. Conditioning on Pre-Treatment GPA				
GPA 2018-II	0.261	0.045	0.504*	0.311
	(0.228)	(0.358)	(0.291)	
% Credits Approved	0.01	0.034	-0.015	0.462
	(0.034)	(0.048)	(0.048)	
Panel C. Conditioning on Pre-Treatment GPA and Freshman Status				
GPA 2018-II	0.291	-0.0528	0.580*	0.162
	(0.222)	(0.358)	(0.298)	
% Credits Approved	0.012	0.026	-0.015	0.545
	(0.034)	(0.049)	(0.049)	
Panel D. Including Region and Faculty Fixed Effects				
GPA 2018-II	0.171	-0.026	0.494	0.220
	(0.230)	(0.347)	(0.310)	
% Credits Approved	0.013	0.019	-0.001	0.742
	(0.034)	(0.049)	(0.049)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column four displays a p-value test of the difference in treatment effect between males and females.

Academic performance outcome variables are GPA, which takes values 0-20, and % of Credits Approved, which is the total credits a students obtains from a given semester divided by the total credits they have applied to complete.

Table 4. Mean Differences in Social Responsibility and Community Engagement

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(122)	(58)	(64)	(Females - Males)
1. I identify with my community	0.126	0.069	0.184	0.754
	(0.187)	(0.235)	(0.290)	
2. I am engaged with the problems of my community	0.444**	0.241	0.618**	0.320
	(0.188)	(0.303)	(0.235)	
3. Usually, I think about solutions for my community's problems	0.313*	0.0690	0.537**	0.185
	(0.178)	(0.254)	(0.252)	
4. Usually, when I am away from my community I feel proud of it	0.228	0.138	0.329	0.571
	(0.175)	(0.208)	(0.271)	
5. Usually, when I am away from my community I am engaged in its maintenance	0.264*	0.241	0.284	0.883
	(0.146)	(0.222)	(0.196)	
6. When I am away from my community, I identify as an ambassador to it	0.132	0.207	0.069	0.731
	(0.205)	(0.274)	(0.302)	
7. I know a depth of my community's problems	0.602***	0.379	0.825***	0.285
	(0.212)	(0.311)	(0.286)	
8. My individual successes are also successes of my community	0.179	0.276	0.075	0.575
	(0.182)	(0.278)	(0.235)	
9. I have the responsibility to help other members of my community get the same opportunities which I have gotten	0.183	-0.310	0.625***	0.0008***
	(0.146)	(0.193)	(0.209)	
10. I believe that I am an example for other members in my community	-0.00942	-0.345	0.292	0.069*
	(0.177)	(0.274)	(0.226)	
11. In general I feel valued by the members of my community	0.240	-0.035	0.478**	0.195
	(0.197)	(0.337)	(0.216)	
12. The problems of my community are also my problems	0.314*	0.172	0.445**	0.445
	(0.179)	(0.287)	(0.221)	
13. I believe that I can be useful to help solve the problems of my community	0.286**	0.0690	0.490**	0.119
	(0.138)	(0.184)	(0.202)	
14. In general, are you willing to assume risks?	0.215	0.241	0.208	0.895
	(0.132)	(0.162)	(0.201)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Mean Differences in Measures of Self-Efficacy and Motivation

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(122)	(58)	(64)	(Females - Males)
1. I can resolve problems if I keep trying	-0.154	-0.379	0.045	0.211
	(0.170)	(0.269)	(0.213)	
2. I always find the way to achieve what I want	0.098	0.069	0.133	0.869
	(0.198)	(0.267)	(0.294)	
3. I never quit my tasks until I have completed them	0.082	0.207	-0.016	0.438
	(0.148)	(0.189)	(0.221)	
4. In general, I confront unforeseen and difficult problems in my life	-0.0358	-0.103	0.031	0.673
	(0.161)	(0.237)	(0.221)	
5. Thanks to my responsive ability, I can manage and resolve unexpected events	-0.014	-0.207	0.165	0.673
	(0.162)	(0.238)	(0.222)	
6. I can resolve the majority of my problems if put in the necessary effort	0.190*	0.138	0.237*	0.607
	(0.098)	(0.135)	(0.142)	
7. If assigned responsibilities, I prefer to distract myself with other tasks so that I have an excuse if I cannot accomplish it	0.165	0.379	-0.019	0.445
	(0.265)	(0.369)	(0.378)	
8. When I have a problem, usually I can find a solution	-0.008	-0.310	0.259	0.07*
	(0.154)	(0.275)	(0.161)	
9. I stay calm during difficult situations because I trust my ability to respond	0.353*	-0.034	0.724***	0.042**
	(0.191)	(0.276)	(0.256)	
10. When I face a problem, usually I can find more than one solution	0.205	0.207	0.214	0.983
	(0.161)	(0.208)	(0.242)	
11. Usually I can face any problem that appears in my life	0.000269	0	0.009	0.974
	(0.151)	(0.244)	(0.190)	
12. When I believe a difficult situation is going to occur, I lose my calm	-0.129	-0.069	-0.161	0.874
	(0.296)	(0.417)	(0.418)	
13. When I have a problem, I prefer to occupy my time focusing on other things instead of addressing the problem	0.267	0.414	0.133	0.599
	(0.269)	(0.396)	(0.369)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Summary Index of Non-Cognitive Outcomes

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(122)	(58)	(64)	(Females - Males)
Social Responsibility Index	0.274**	0.000	0.525***	0.034**
	(0.126)	(0.177)	(0.178)	
Community Engagement Index	0.425***	0.253	0.591***	0.251
	(0.149)	(0.199)	(0.222)	
Sense of Belonging Index	0.202	0.138	0.255	0.661
	(0.133)	(0.216)	(0.165)	
Empowerment Index	0.138	0.0862	0.194	0.656
	(0.123)	(0.167)	(0.180)	
Self-Efficacy Index	0.0526	-0.0805	0.175	0.227
	(0.105)	(0.168)	(0.133)	
Confidence Index	0.0709	0.0345	0.117	0.773
	(0.145)	(0.213)	(0.197)	
Motivation Index	0.175	0.310	0.0588	0.454
	(0.170)	(0.249)	(0.231)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Social responsibility index is created by Table 4 statements 3, 9, 12, 13. Community engagement index is created by Table 4statements 2, 4, 7.

Sense of belonging index is created by Table 4 statements 1, 5, 8, 11. Empowerment index is created by Table4 statements 6, 10, 14.

Self-Efficacy index is created by Table 5 statements 1, 5, 6, 8, 10, 26.

Confidence index is created by Table 5statements 4, 7, 9, 11, 12. Motivation index is created by Table 5 statements 3, 13.

Table 7. Falsification Test

	I	II	III	IV
	Full Sample (131)	**Males Only** (64)	**Females Only** (67)	**P-value Test** (Females - Males)
Panel A. Simple OLS				
GPA 2017-I	-0.119	-0.715	0.477	0.07*
	(0.339)	(0.480)	(0.463)	
% Credits Approved	-0.0002	-0.026	0.025	0.429
	(0.033)	(0.042)	(0.05)	
Panel B. Conditioning on Freshman Status				
GPA 2017-I	-0.213	-0.444	0.140	0.322
	(0.307)	(0.423)	(0.427)	
% Credits Approved	-0.006	-0.013	-0.004	0.884
	(0.032)	(0.041)	(0.05)	
Panel C. Including Regional and Faculty Fixed Effects				
GPA 2017-I	-0.191	-0.358	0.128	0.421
	(0.316)	(0.432)	(0.499)	
% Credits Approved	0.006	-0.004	0.023	0.613
	(0.028)	(0.035)	(0.046)	

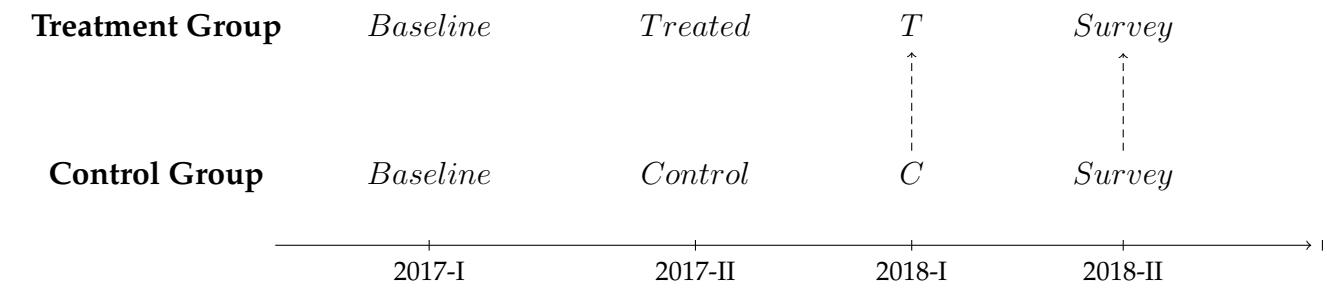
Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column four displays a p-value test of the difference in treatment effect between males and females.

Academic performance outcome variables are GPA, which takes values 0-20, and % of Credits Approved, which is the total credits a students obtains from a given semester divided by the total credits they have applied to complete.

7 Figures

Figure 1. Timeline of Randomization



Where period 2017-I is the baseline period before treatment, 2017-II is when the treatment period commences, 2018-I when the treatment period ends and our analysis begins. Finally, period 2018-II is when an exit survey is conducted and final comparisons are made.

8 Appendix

Appendix Table 1. Difference in Differences, Academic Performance

	I	II	III	IV
	Full Sample	**Males Only**	**Females Only**	**P-value Test**
	(129)	(64)	(65)	(Females - Males)
Panel A. 2018-I Academic Performance				
GPA 2018-I	0.423	0.061	0.660	0.481
	(0.440)	(0.645)	(0.576)	
% Credits Approved	0.054	-0.018	0.130*	0.136
	(0.05)	(0.069)	(0.073)	
Panel B. 2018-II Academic Performance				
GPA	0.430	0.133	0.604	0.564
	(0.423)	(0.615)	(0.563)	
Credits Approved	0.052*	0.047	0.008	0.872
	(0.028)	(0.042)	(0.073)	

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each panel controls for a student's freshman status in a given semester.

Column four displays a p-value test of the difference in treatment effect between males and females.

Appendix Table 2. OLS Regression of 2018-I Treatment Effects Controlling for Monetary Reimbursements

	I	II	III
	Full Sample	**Males Only**	**Females Only**
	(129)	(64)	(65)
Panel A- Simple OLS, Controlling for Monetary Incentive			
GPA 2018-I			
Treated=1	0.356	-0.193	0.904
	(0.730)	(0.942)	(1.087)
Amount Received	-0.001	-0.001	-0.0007
	(0.004)	(0.005)	(0.006)
% Credits Approved 2018-I			
Treated=1	0.0204	-0.0290	0.049
	(0.085)	(0.107)	(0.137)
Amount Received	0.0002	-0.00001	0.0005
	(0.0005)	(0.0006)	(0.0007)
Panel B- Including Regional Fixed Effects			
GPA 2018-I			
Treated=1	-0.362	-1.074	0.758
	(0.722)	(0.865)	(1.221)
Amount Received	0.004	0.0056	-0.0003
	(0.004)	(0.005)	(0.007)
% Credits Approved 2018-I			
Treated=1	-0.039	-0.081	0.024
	(0.104)	(0.102)	(0.204)
Amount Received	0.0006	0.0004	0.000871
	(0.0006)	(0.0006)	(0.001)

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each panel controls for a student's freshman status in a given semester.

Column four displays a p-value test of the difference in treatment effect between males and females.

Appendix Table 3. Monetary Reimbursements on Academic Performance, A Correlation Test Among Treated Students

	I	II	III
	Full Sample	**Males Only**	**Females Only**
	(129)	(64)	(65)
Panel A. 2018-I Academic Performance			
GPA 2018-I	-0.00103	-0.00140	-0.000705
	(0.00426)	(0.00544)	(0.00628)
% Credits Approved	0.000192	-0.0000139	0.000510
	(0.000508)	(0.000642)	(0.000787)
Panel B. 2018-II Academic Performance			
GPA 2018-II	-0.000762	-0.00330	0.00207
	(0.00394)	(0.00531)	(0.00617)
% Credits Approved	-0.000452	-0.000865	0.000255
	(0.000463)	(0.000658)	(0.000648)

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Each panel controls for a student's freshman status in a given semester.

Column four displays a p-value test of the difference in treatment effect between males and females.