The Unexpected Effects of a Merit-Based Scholarship: Evidence from Ser Pilo Paga

Nathalie Basto Aguirre

Advisers: Prof. Juan F. Vargas and Prof. Jorge A. Gallego
Universidad del Rosario
Economics Department
Master of Science in Economics
Bogotá, Colombia
June 2019

The Unexpected Effects of a Merit-Based Scholarship: Evidence from Ser Pilo Paga

Nathalie Basto Aguirre*

Bogotá, Colombia
June 2019

Abstract

This article examines the unexpected impacts of the need-and-merit-based college scholarship program Ser Pilo Paga (SPP) on schools aggregated academic performance and dropout rate. My analysis uses difference in differences methodology to estimate variation in average test scores across cohorts from schools that had students selected for SPP. SPP could have had an impact on schools' results through different potential mechanisms. SPP could emerge motivational positive incentives to increase effort, related with better academic results. This incentive seems to be stronger for last year students. SPP could also cause the reduction of dropout rates caused by the incentive scheme to keep studying until the last year of school and have the opportunity to attend university, producing a change in the composition of the distribution of students' skills. Moreover, schools may react strategically in order to meet students', parents' and public demands to have more selected students. My results suggest that schools that had SPP students performed better in standardized test for 11th grade, but worse in 9th and 5th grade standardized tests. After dealing with some issues related to parallel trend assumption, my estimates for high school dropout rates suggest that SPP reduced such rates supporting the change in composition hypothesis. I also analyze heterogeneous effects on the proportion of need-eligible students and I find that the effect of SPP on outcomes becomes more positive in schools that have a bigger proportion of eligible students.

JEL classification: I20, I28.
Keywords: Incentives; Student achievement; scholarship; schools behaviour.

*Universidad del Rosario. Correo-e: nathalie.basto@urosario.edu.co

1 Introduction

The lack of opportunities for social mobility, such as limited access to higher education, can discourage human capital accumulation, especially for the poor, contributing to the endurance of poverty, inequality and affecting economic growth (Kaldor, 1957; Stiglitz, 1969; Okun, 1976). Therefore, governments have made efforts to ensure inclusive access through the implementation of different types of public programs. Based on the relationship between socioeconomic background and education enrollment (McPherson and Schapiro, 2006), initiatives that provide better access opportunities for students from lower income families are widespread. Among these initiatives, scholarships and student loans are the most common ones.

Scholarships based on academic merit may act as high-powered incentive schemes and should align the behavior of agents in the system, as in the case of public policy in education design. Nevertheless, incentive schemes may produce distortions in the expected results and behavior, especially along other dimensions and agents, partially related with the dynamics of the scholarship program. Literature has shown evidence of unintended effects of incentive schemes when they involve measuring performance (Baker, 1992), dealing with multitasking skills of the actors involved (Holmstrom and Milgrom, 1991) and depending on specific tasks (Glaeser and Shleifer 2001). Financial aids should act as incentives promoting effort and academic performance (Lazear, 2000) and are expected to generate better academic results among the target population. However, there is empirical evidence of unintended effects of education policies when the incentive schemes change, not exclusively for the target population or direct beneficiaries (Jacob and Levitt, 2003; Jacob 2002).

Ser Pilo Paga (SPP), a need-and-merit-based college scholarship program in Colombia, offers a good framework to explore unexpected impacts of merit-based financial aid policies not only on direct beneficiaries but also on potential beneficiaries, school students and schools, by analysing their academic achievement and attainment outcomes. The program was launched at the end of 2014 school year and it ended in 2018. SPP offered a particularly good opportunity to access higher education for best performing students with low financial resources to attend college, in a country where less than 25% of the 25-64 years old population have attained tertiary education (OECD, 2016). The eligibility conditions for SPP were to be above a certain cut-off value in the 11th grade standardized test (Saber 11) and below a certain cut-off in SISBEN[1]. Since the program had a considerable scale and relevance for the public opinion in the country, and the academic merit condition is determined using scores from the national standardized exit examination, Saber 11, it could prompt motivational efforts to emerge among 11th grade students and in lower grades, creating new incentives for the school community as well.

I analyse whether the new opportunities, generated by SPP, to access higher education had unexpected effects on schools academic and achievement results. This article evaluates the effect of SPP on standardized tests

[1] Index used to measure household wealth

for 5th, 9th, and 11th grades and also explores the impacts of the program on dropout rates using school level panel data. This study identifies the effect of SPP using a difference in differences approach relying on a comparison of academic outcomes across schools that had at least one SPP student in the first year of the program and those that did not. I primarily consider average effects but explore heterogeneous effects depending on the proportion of need-based eligible students.

My results show a moderate positive effect on 11th grade test scores that is especially robust for math scores, but negative effects of SPP on 9th and 5th grade tests scores in both math and language. Regarding dropout rates, my estimates suggest a positive effect of the program in reducing such rates, supporting the hypothesis of a possible change in the composition of student's skills leading to worse results in lower grades' tests scores. These results can also be explained theoretically by a student motivational effect that only emerges for high-skilled senior students and by a concentration of school efforts to maximize the number of SPP students, focusing on 11th grade need-eligible, high skilled students.

The main contribution of this paper is to extend the understanding about scholarships and financial aid mechanisms and their incentive schemes on school outcomes, by exploring the effects of this intervention on potential beneficiaries rather than on direct beneficiaries. My results propose important lessons for policy makers, as scholarship programs may act as sufficient incentives to change dropout decisions. The education system should guarantee an incentive framework to increase quality in education for those students who are staying in school. There is a need to provide stronger incentives for schools and students to raise performance not only in senior year but also along all school years to develop long lasting skills.

The remainder of the article proceeds as follows. Section 2 explains the SPP program and its context in Colombia. Section 3 makes a literature review on the potential effects of the program on schools outcomes and its relation with the potential effect of SPP. Section 4 presents the data. Section 5 explains the empirical strategy. Section 6 discusses the results and section 7 presents a final discussion.

2 Context of Ser Pilo Paga (SPP)

In recent years, Colombia has made efforts through different public policies to improve the quality of basic education and extend the access to higher education (DNP, 2015). Policy has also been oriented towards increasing education achievement outcomes resulting from the persistence of high inequalities in students' academic results among schools. Such results are partially explained by the segregation of Colombian schools according to students' socioeconomic status, being detrimental for schools with poorer students (Duarte et al. 2012). This becomes more relevant as we consider that several studies have found that the school characteristics strongly impact academic performance in developing countries (Gamoran and Long, 2006; Banerjee et al. 2007; and Behrman, 2010).

Despite major achievements in terms of coverage in primary and secondary education, Colombia still faces vital challenges in education quality and also in ensuring school completion and access to tertiary education (DNP, 2015) (OECD, 2016). The education net enrollment rate for primary and secondary education was 84.9% in 2017, and when focusing on upper secondary education it drops dramatically to 42.79%[2] . The government and organizations have identified a high dropout rate in school, 3.08% in 2017, partially explained by the dropout rate during upper secondary education, 3.94%.

Access to higher education is limited for a great number of students, because most universities in Colombia are private and expensive (OECD and The World Bank, 2012). While private institutions exhibit higher returns (Camacho, Messina and Uribe, 2016; MacLeod, Riehl, Saavedra and Urquiola, 2017), their tuition fees are more than double their public counterparts (Londoño et al., 2018), which are subsidized by the government but have a very limited capacity to receive new students. Private universities are a privilege denied to many by the tuition fees they charge (Riehl et al., 2016) along with few options of external financial aid and little access to credit markets for students loans (OECD and The World Bank, 2012; Velasco and Sanchez, 2014). This leaves most low-income students without the capacity to afford attendance costs at postsecondary institutions, especially universities.

In this context, the Ministry of Education created Ser Pilo Paga (SPP). This program offered a loan/scholarship to pursue undergrad studies in high-quality accredited universities (both private and public), and it became a scholarship upon graduation. The beneficiaries also received a living stipend depending on the city they are studying and where they come from.

SPP eligibility conditions were based on academic performance and a need-based criterion. First, it was required to be above the score cut-off on the Saber 11 test, the standardized exit examination in Colombia. The government increased it every year after students had taken Saber 11 test to adjust the number of potential beneficiaries to the budgetary restrictions of the Ministry of Education. Second, it was necessary to have a SISBEN score (the index used to measure household wealth and to target social programs as a proxy to poverty) below a cutoff point also defined by the government. Finally, students should be accepted in a high-quality accredited University. Among these criteria the hardest one to accomplish is to be above the cutoff score on Saber 11. Only 10% of test takers achieved this result, compared with 55% of students whose SISBEN score is below the cut-off in 2015 (Laajaj et al. 2018).

SPP was launched on October 1st, 2014, just two months before the end of the school year, and after 11 graders from class of 2014 had taken Saber 11[3] . The first cohort of students that could have been beneficiaries were high school graduates in the second semester of 2014 who began studying at the university

[2] Source: MEN-SIMAT / proyecciones de población del DANE ajustadas censo 2005. Ministerio de educación nacional. Oficina Asesora de Planeación y Finanzas. COLOMBIA: Cobertura Bruta y Neta por Entidad Territorial.

[3] Then in this first cohort, students were not aware that they have to obtain a certain minimum result in the test to have access to a scholarship or even the potential existence of the program.

in the first semester of 2015[4]. Annually, the program awarded an average of 10,000 scholarships reaching nearly 40,000 beneficiaries in 2018 with its fourth and final cohort. For this paper, I use the data of the three first cohorts of the program gathering 31,940 beneficiaries. In Table 1. I present a summary of the eligibility conditions.

Table 1: SSP elegibility conditions

VERSION	Eligibility conditions					
	Saber 11		Maximum SISBEN score			No. of high quality accredited universities in the country
	Minimum score	Date of the test	Metropolitan areas	Other urban areas	Rural areas	
SPP1 (began to study in 2015)	310/500	03-agu-14	57,21	56,32	40,75	33
SPP2 (began to study in 2016)	318/500	02-agu-15	57,21	56,32	40,75	39
SPP3 (began to study in 2017)	342/500	31-jul-16	57,21	56,32	40,75	44

SPP was a well-recognized national scale and controversial program in Colombia. It had the potential to transform the opportunities for selected students. However, it was perceived as problematic by some sectors as it would take public resources that could potentially be redistributed to public universities increasing the access to higher education and offering more places at subsidized fees.

SPP also changed the perceptions of students and schools about access to higher education. A short-term impact evaluation of the program (Alvarez et al, 2017) found that the probability of going to the university is between 25% and 32% higher for SPP beneficiaries; SPP also generated an increase in the possibility to attend better-quality universities and reduce university first semester dropout rates (Londoño et al., 2018). Moreover, in the official impact evaluation, performed with group interviews and focus groups among selected and not selected students close to the cutoff values of Saber 11, CNC and Universidad de los Andes (2016) found that the perception of the opportunities to go to college was very limited before the program. They found that students and their families' expectations about their future increased after the program and that eligible students reported having friends or siblings who have already been beneficiaries, in line with a potential peer effect.

Besides, there might have been an increase in Saber 11 test preparation. Interviewed students considered that receiving school preparation for taking the exam could have been a crucial factor in obtaining the required score (CNC and Universidad de los Andes, 2016). From questionnaires conducted to administrative personal at test-taking preparation academies[5], I collect anecdotal evidence, finding that students have increased their

[4] Because of Colombia's academic year.
[5] Audios available upon request.

enrollment in courses that guarantee the minimum score required by SPP, and decreased in the ones that do not. Academies' staff partially explained this claiming that since the beginning of the program, schools have created their own extracurricular preparation on test-taking for Saber 11. This evidences that the program might have also changed schools' dynamics.

Additionally, SPP raised public attention on school performance. Since SPP was a nationwide program that gives a lot of public recognition to selected students and their schools[6], it could have generated a change in schools' behavior. From conversations with policy makers, I found anecdotal evidence that parents started to think of the importance of attending a school that has better results in Saber 11, because it increased the possibility for their children to become SPP beneficiaries. Many schools which obtained a large number of beneficiaries were recognized by the local and even the national government in public events. Schools with beneficiaries also caught the attention of local media, with school principals being interviewed about the strategies they have implemented to obtain better Saber 11 results. In Appendix I, there is a collection of news and online articles that present evidence of what is mentioned above.

In sum, It is interesting to study the effect of the program not only on the beneficiaries, university students, but also on the potential impact it could have had on the behavior of school students, both younger and senior, and schools. These effects should be seen at the schools aggregated results. Therefore, in an attempt to deepen in the understanding of unexpected effects of this scholarship program on academic achievement and attainment, I will assess changes at the aggregated school level not only for 11th grade tests scores but also for the standardized tests performed in 5th and 9th grade. Moreover, I will look for the effect of SPP on dropout rates, to evaluate whether incentives are strong enough to generate anticipated changes in effort and dropout decisions among younger students and to understand the average effect of the program on schools results. Now that the program is over, it becomes even more critical to learn lessons about SPP in order to improve the design on future public policies that subsidize the demand for higher education.

3 Related Literature

Literature has widely studied the effects of financial aid for students (Cohodes and Goodman, 2014; Scott-Clayton and Zafar, 2016; Solis, 2017). A broad empirical literature review concludes that the effect of financial aid or incentives, especially on academic performance, is still under discussion[7]. The effects of SPP on the beneficiaries have also been studied. SPP increased immediate college enrollment more than

[6] Check Appendix I for news related to public recognition by media and local authorities to schools with SPP selected students.

[7] Such evidence ranges from studies which find no effects (Angrist and Lavy 2002), negative effects (Cornwell et al. 2003), positive effects on achievement of high-ability students, but negative impact on achievement of low-ability students(Leuven et al. 2003) (Levitt et al., 2016) to research that finds positive and large effects in academic performance and completion (Bettinger, 2004; Angrist and Lavy, 2009; Barrera-Osorio, et al. 2011; Scott-Clayton 2011; Fack and Grenet 2015). To look for a wider discussion on financial incentives on academic performance (Deming and Dynarski, 2010; Cornwell, Mustard, and Sridhar, 2006 ;Belot et al. 2007; Angrist et al. 2009; Mealli and Rampichini, 2012; Solis 2017; Banerjee et al. 2017)

80% driven by matriculation at private high-quality colleges, closing the socioeconomic enrollment gap and increasing diversity by 46% (Londoño-Velez et al., 2018).

While most of these studies have focused on analyzing the effects of the program on university students and direct beneficiaries but not ex-ante effects of the program, there is a growing interest in incentives to enhance school students' academic results as a cost effective policy, changing schools and school students' incentives and behavior.

3.1 Potential effects on students

Financial incentives or scholarships can generate a motivational effect to choose higher levels of effort that correlates with high tests scores and can potentially motivate greater investment in education by students and their parents increasing graduation rates (Lazear,2000; Barrow et al., 2014; Bettinger, 2012; Fryer, 2011). In the case of SPP, by increasing the probability of going to college subject to performance, one may expect that students choose higher levels of effort correlated with knowledge and skills gains which will then translate into better performance in tests. However, critics sustain that these gains can be related only with test taking skills (Jacob, 2002) or that these types of incentives could even decrease the intrinsic motivation of students (Deci 1972a, 1972b; Kruglanski, Friedman, & Zeevi, 1971).

According to the model and empirical evidence presented by Levitt et al (2016), De Paola, Scoppa and Nistico (2012), and Leuven et al. (2003; 2010) motivational incentives are stronger in students whose marginal return of increasing their effort is high enough. For instance, in the case of Saber 11 test scores, students whose distribution of their expected test scores is centered at the merit-based eligibility threshold or just below it should be the most motivated by the introduction of SPP (Laajaj et al., 2018). In other words, high ability students who are aware of their capability of achieving the expected result if they increase their effort are more likely to have a positive effect for the program. In contrast, low ability students need to make great amounts of extra effort with lower probabilities of getting the desired result; then their marginal return of effort is lower and are less likely to change their effort decisions.

In the same line, SPP incentives should be stronger for need-eligible students, who have a higher expected return on increasing efforts and for those students whose school had an SPP student in previous years. By perceiving a higher probability of getting the scholarship conditioned by their effort and having a closer contact with the benefits of the program, they may see the program as a real opportunity to change their future.

Following the argument proposed in Levitt et al. (2016), SPP may increase the expected opportunity cost in dropout decisions for the one, marginal student who is just indifferent between continuing in school and dropping out. This potential effect focuses on those who are indifferent and others whose gain in expected return of completing high school caused by the program is enough to change their opportunity cost, consequently having greater benefits to keep studying than to drop out.

However, theory also suggests that long-run motivational incentives are weaker for younger students, who are more likely to react to immediate than future rewards (Angier, 2009). Students in lower grades may perceive the saber 11 test or the election for the scholarship as a long-term event. As teenagers, they might take riskier choices (Casey et al., 2008; Dahl, 2004; Ernst et al., 2006; Spear, 2009), not altering their behavior with anticipation. Consequently, in this scenario and assuming no response in schools to the program, one would not expect to see significant increases in Saber 5 and Saber 9 test results caused by a motivational effect of the program.

Nevertheless, to have the chance to go to college and obtain the scholarship, students must complete high school, so incentives to change dropout decisions should be stronger than those to change effort decisions. As low academic performance is one of the most recurrent factors related to dropping out (Lan & Lanthier, 2003), one would expect that student dropout decisions affected by the program are more likely to happen among low skill/ low performer students, that now will continue studying. This would change the distribution of students' skills for certain schools and certain grades, potentially affecting aggregated academic performance. Those incentives are fostered for students whose schools had SPP alumni in previous years, where one would expect better academic quality and increases in parents expectations for children's education, both positively related to the likelihood that students would graduate from high school (Fetler, 1989; Rumberger, 1995).

Regarding students incentives, peers and older classmates are usually seen as reference points (Brickman and Campbell, 1971; Kahneman, 1979, 1999; Frederick and Loewenstein, 1999; Lin, 2010). This motivational effect may also be reinforced when students perceive that their schoolmates have the opportunity of being selected for the program, go to college or merely perform better (Schreck et al., 2014; Sacerdote, 2011; Albert et al., 2013). Literature related with peer effect on students' outcomes also suggests that such motivational effect could be present even for non-need-eligible students. Having classmates selected for academic merit programs or rewards can have an impact, for instance, in academic performance, college enrollment and dropout decisions (Sokatch, 2006; Tokuno & Campbell, 1992; Cornwell et al., 2003, 2006).

3.2 Potential effects on schools

Incentive theory also suggests possible effects of incentive schemes based on academic performance on school behavior. SPP acting as an incentive policy may have created some sort of non-consequential accountability system as proposed by Hanushek and Raymond, 2005, which does not have direct consequences on students or school staff. Nevertheless, the public attention generated by the program, especially for those schools that had SPP students in the initial year of the program, could have increased pressure of students and parents to get better results in Saber 11. This pressure for accountability may cause schools to take better teaching strategies, improve curriculum and incentivize teachers' efforts benefiting all students (Jacob, 2002).

However, there is also evidence of incentives on performance causing teachers to boost their students' grades to increase their access to college and to meet the students and parents demands (Bugler, Henry and Rubenstein, 1999). Also, test results manipulation may occur through excluding some students from testing or increasing grade retention, so that low performer students do not affect average test results (Klein et al. 2000, Jacob, 2002; Jacob and Levitt, 2003; Cullen and Reback 2002; Figlio and Getzler 2002).

Additionally, multitask agency models imply that incentives to perform on one dimension may cause stakeholders to remove efforts from other dimensions. For instance, teachers may spend less time on untested material when the rewards are concentrated on test performance, or students take less difficult courses when they are rewarded for having high grade point averages (Binder, Ganderton, and Hutchens, 2002; Cornwell et al. 2006; Glewwe, Ilias, Kremer, 2003).

Schools that had at least one selected student in previous years should be the ones that exhibit stronger incentives to change their behaviour. Mainly, because they were in the eye of the public attention facing bigger pressure or incentives to perform better. Similar to Levitt et al. (2016), schools with at least one SPP student know that they are capable of achieving such results and then have higher incentives to change their behavior, in comparison with schools that did not have any SPP students, especially if the highest score obtained by one of their students was way below the cutoff value.

The only previous study that assesses unexpected effects of SPP on school students is Laajaj et al. (2018). They use a difference in differences approach and a regression discontinuity design to look for differences in Saber 11 tests scores between need-based eligible and non-eligible students. Laajaj et al. find that need-based eligible students performed better in Saber 11 if they are in the top 25% of the performance distribution. This reinforces the idea that the stronger incentives are for those need-eligible high skill students for whom the program significantly increases the expected marginal return of making additional effort. They argue that SPP increases the incentives for low-income high school high achievers and improves academic performance, advocating for the virtues of the SPP program. However, when analyzing aggregated school outcomes for saber 11, Saber 9 and Saber 5, it is not obvious to expect a positive effect of SPP, assuming that across school and year the proportion of students for whom the program rises significantly their expected return of additional effort is smaller than those for whom the motivational incentives are not strong enough.

Therefore, it is difficult to conclude that the program might have an impact on learning and knowledge outcomes or on academic performance by only studying standardized test results of 11th grade from the previous year. Hirshleifer (2017) states that learning production processes require sustained effort over time, but according to Kaur, Kremer, & Mullainathan (2015) agents may lack the self-control to exert effort now in order to earn rewards later. While it is not possible for me to test directly whether the gains in Saber 11 scores are due to "Studying for the test" and therefore just based on test-taking skills, since I have no data from alternative tests or measures of skills for 11th grade students. However, it is worth looking for the impacts of the program in a broader set of outcomes that provides evidence of larger impacts of SPP on academic performance.

The discussion presented in this literature review highlights the possibility that the program creates an incentive scheme that has an effect in both students and schools' behaviour. This scheme should also affect need-based non-eligible students and students in lower grades, standing out the importance to explore the effects of the program in a broader set of outcomes. In an attempt to disentangle the unexpected effects of the program, the present article assesses effects of SPP on aggregated academic school results, not only for 11th grade, but also for 5th and 9th grade. This contributes to understanding the strength of the motivational incentives generated by the program and provide ample evidence regarding the impact of the program on students effort, gains in knowledge and lasting skills, in opposition to effects only in the test that define the academic criterion for the SPP program. Moreover, I analyse the effects of the program on dropout rate. The literature suggests that the SPP effect could be positive in having students finish school, but at the same time, it could have a negative impact in test results by changing the composition of students skills as lower skill students choose to remain in school.

The data used in this article, described in detail in the next section, does not allow me to identify separately the effect of the program on students and schools in aggregated educational outcomes. Hence, my estimations should be interpreted as the combined effect of the incentives generated by SPP on students and schools. Nevertheless, I provide evidence supporting the hypothesis that there is a motivational incentive that reduce dropout rates and could affect the composition of skills inside school and grades. Moreover, aggregated school panel data allow me to look for causal impacts of the program on schools' Saber 11, Saber 5, and Saber 9 tests scores and school dropout rate.

4 Data

This paper uses two main sources of administrative data. First, I use the R166-SIMAT database from 2012-2017, which contains census information from all enrolled students in public schools in the country at the level of school. Second, I use Saber databases to obtain Saber 5, 9, and 11 at the school level. Finally, I use information from SISBEN, processed by the Ministry of Education from 2012 to 2016 to look for heterogeneous effects. In the following subsections, I will describe this information and characterize the sources of variation I will use in the estimations.

4.1 Census of students in public schools at the school level

The Ministry of Education started collecting information from all students enrolled in the country in 2005. This database, known as the R166-SIMAT[8]. For the purpose of this paper, I will use the database from 2012 to 2017 and only including students between the first grades of primary education to the last grade of upper secondary education, 11th grade, at the level of school. It allows me to know the dropout rate for every school and every grade, since the effect of the existence of the program and having a schoolmate selected for the program could have different impacts for different grades depending on how close they are to the last year of school. This database also allows me to have socioeconomic variables from school such as the average socioeconomic strata of their students, the proportion of women and victims of the conflict, whether the school is public or not, whether the school is located in a rural or urban area, among others.

The abovementioned data excludes schools that have nontraditional school models, for instance certain flexible education models for some indigenous communities in rural areas; it also excludes public institutions for adult education and literacy, and "escuelas normales" (teacher training colleges). In addition, I will use the data for the schools where at least one of their students took Saber 11 for each year, meaning that they have a Saber 11 test score. Considering that I am only using schools that have Saber 11 score, it excludes schools that have only primary grades or primary and lower secondary grades, or schools that in a specific year did not have any 11th grade student. In the end, I keep about 50% of the schools in SIMAT during this period. After this process I have 8,701 schools per year that receive approximately 7.3 million students, and nearly 71% of them are public.

Based on official information from the Ministry of Education, which uses only students in public schools, dropout rate decreased from 4.9% in 2010 to 3.1% in 2017 mostly based on a reduction in dropout rate in primary education.

4.2 Saber – ICFES

These datasets contain the score reports in the standardized official exit examination, reported by the Colombian Institute for Higher Education Promotion - ICFES (acronym in Spanish). ICFES assesses students and school performance through standardized testing in 5th, 9th, and 11th grades; this national tests are called Saber 5, Saber 9, and Saber 11 respectively. It also collects other socioeconomic characteristics. I use the information of scores in Saber 11 (S11) that contains information of about 550,000 test takers per year, and scores of Saber 5 (S5) and Saber 9 (S9) at the school level. ICFES does not process the information for the latter two exams at the individual level and reports the aggregated scores only for schools whose number of

[8] In 2004, the Ministry of Education decreed the "Resolución 166", which mandated that all education institutions must report, amongst other information, all their students enrolled individually. The R166 is currently known as SIMAT (National Enrollment System, as per its Spanish acronym).

test takers is big enough to comply with their statistical requirements.

Saber 11 has a global score for the test since 2014, but before that, it reported the scores for different subject areas separately, these areas have suffered different changes and regroupings through the years. However, Math, or quantitative reasoning test, has always been conducted, as well as (reading) Language[9]. So, I decide to focus on those components, as they are often used on several papers about education achievement regarding data from Colombia (Bettinger, 2005; Baez and Camacho, 2011; Guarín et al. 2016). For S5 and S9 the only separate report at the school level scores are for Math and Language[10]. In Table A.I. in the Appendix, I present summary statistics for tests scores in each year for the universe of schools that have information for all tests scores measures. There is an increase in average Saber 11 tests scores for both math and language in the period 2012 – 2017, that is bigger for language, similar to the behavior of Saber 5 scores, although the former presented a lower growth. The growth for S9 is lower, and in this case is led by math score.

In order to have comparable schools, I work only with schools that have information on S5, S9, and S11; this is the complete panel for the education achievement outcomes. This excludes those schools that have only secondary and upper secondary education but not primary and those that do not have S5 and S9 tests scores reported by ICFES, which corresponds to 15% of all previous observations[11] . To be part of the panel, a school must have information from both SABER and C600 for a given year, but it does not necessarily have to have information for the six periods. Then, I work with an unbalanced panel.

ICFES is responsible of reporting to the Ministry of Education and ICETEX about those students who have been selected as beneficiaries, according to both academic (S11 results) and need-based criteria (SISBEN). That information is also reported to schools and students for them to start the application processes. I have the database with the annual number of SPP students in each school from 2014, when the program started, to 2017. In Table 2. I present the number of schools that have at least one student selected for the program. The percentage of these schools starts at around 42.53 eventually reaching 48.54, probably because of raised awareness about the existence of the program, changes in students, or schools attempting to increase the number of beneficiaries. In 2016, it falls to 36.88 as the academic criteria become harder, following the adjustment the Ministry of Education performed as a result of the budget assigned to the program.

[9] The general methodology and results is presented in http://tiny.cc/3u8e7y

[10] The general methodology and results is presented in http://tiny.cc/z58e7y

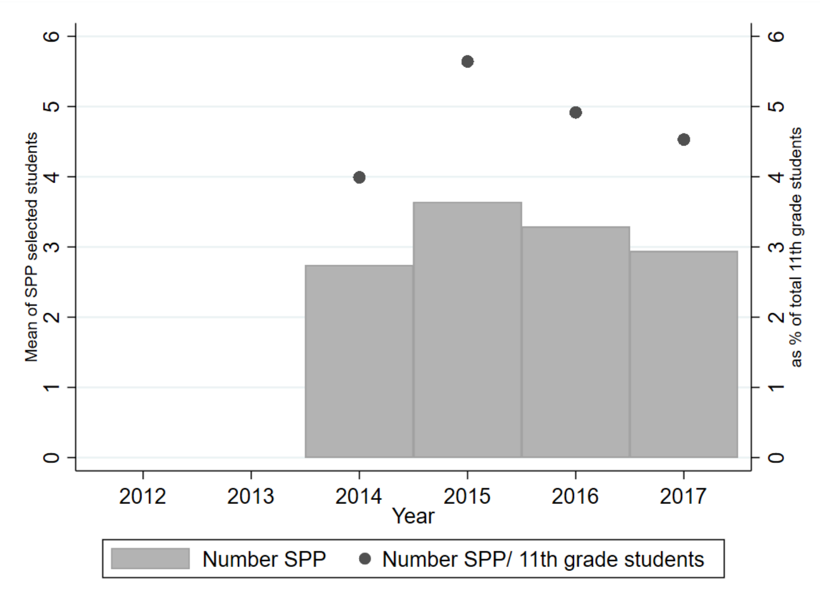
[11] Unfortunately, some school identification codes could not be matched across time because some small schools were merged into larger "education institutions" with a single administration, and others may have been divided into smaller schools dedicated exclusively to primary or secondary education, changing their identification codes without traceability from the Ministry of Education.

Table 2: Percentage of schools with at least 1 beneficiary

Year	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**Total**
#SPP>=1	0	0	42.53	48.54	36.88	28.3	27.96

The average number of SPP students can be seen in Figure 1: between 2.7 and 3.6 students per school were selected during the years of the program following the same tendency explained above. Figure 1 also shows that in schools with at least one beneficiary, between 3.9% and 5.8% of their 11th grade students meet the requirements for SPP.

Figure 1: Average number of SPP selected students
In schools with at least one selected student



5 Empirical Design

As mentioned before, this paper estimates the effect of the incentives generated by SPP on selected educational achievement outcomes, standardized Saber 11, Saber 9 and Saber 5 scores, and on the dropout rate. To analyze the effect SPP has in schools' educational achievement outcomes, the main difficulty is the national coverage of the program given the absence of a natural control group.

This methodology relies on the assumption that schools with at least one SPP student had stronger incentives to change their behavior in the same line that students having older classmates selected for the program have stronger incentives to change their decisions. For some students, meeting the academic criterion to be accepted in SPP could have seem easy while for others, even with similar academic performance, it could have seem very difficult. The empirical strategy assumes that there are important differences in the perception of the probability of having selected students or being selected for the program. Thus, the incentives generated by the program will have a greater impact when expected marginal effort is relatively more effective in the production function. That is, the treatment effect is seen for whom exerting additional effort has the highest expected marginal return, in terms of the probability of meeting the achievement standard and receiving the reward.

This assumption considers the discussion presented in the related literature section. The existence of the program from 2015 onwards is not enough to change the incentives of schools and students. They need to perceive that it is actually possible to be strong enough (Locke & Latham, 1990). In this case schools and students need to perceive that it is possible to have selected students or being selected conditional on attending that specific school . This is stronger when the school have had SPP selected students. Additionally, in schools that have had a selected student, the aggregate effect on students or school staff is more probable to be bigger than for schools that have never had any beneficiaries, in line with the peer effect literature and the literature of how incentive and accountability schemes on results can change schools' behavior. Schools that had a beneficiary in a previous cohort could face bigger pressures to take actions that increase the number of selected students next year.

Therefore, I use the variation on schools' exposition to selected students, comparing the education achievement results of schools with at least one beneficiary in the first year of the program with the results of schools with no beneficiaries, before and after the existence of the program. Using the panel data at the level of school I implement a difference-in-differences with multiple time periods approach using as a control group the schools that did not have an SPP student in previous years, in line with the two-way, fixed-effects, regression models discussed by Bertrand et al. (2004), Wooldridge (2005), Belloni et al. (2014), Borusyak and Jaravel (2017).

The first year in which a school got their first cohort of selected students in SPP varies among schools between 2014, 2015 and 2016, with a high correlation between having selected students in 2014 and having selected students in the next years. To avoid potential endogeneity and to have a better control group for each year in my sample, I define only those schools that had an SPP student in 2014 as treated. In addition, only schools that have never had an SPP student are defined as never treated schools, and constitute my control group. That is to say, I exclude schools that had their first elected SPP in 2015 or 2016 which correspond to approximately 25% of the schools in my sample. However, in Appendix III. I perform a difference in differences approximation with multiple treatment timing, that take into account those schools that had their first SPP student in different years than the first year of the program, finding similar results that are presented in detail in the Appendix.

Moreover, to refine the control group and make it more comparable, I use as a control those schools that could have had selected students in 2015 if the government had not changed the merit-based criterion of 2014 and that could have had selected students in 2016 if the government had not changed the merit-based criterion of 2015. In other words, I use as a control those schools that in 2015 had at least one student whose score in Saber 11 was above 310 and above 318 in 2016, but that have never had SPP selected students. In Appendix II. I present the distribution between treated and never treated schools in each year, histograms of the distribution of SPP selected students in 2014 for the treated group and the whole sample, and a summary of the descriptive statistics for the main variables, for both the treatment and control group before 2015.

In Appendix IV, I present these results when using the same criterion for refining control group, but only for 2015, which lead to the same conclusions. Moreover, I performed the same exercise without refining the control group, that is, using the whole group of never treated schools as control[12]. I find similar robust results for all but one outcome variable, saber 11 math score and only when using alternative definitions of the treatment explained below.

Specifically, I estimate a model of the form:

$$Y_{it} = \beta_1 D_{i2015} + \alpha X_{it} + \pi_t + \pi_i + \varepsilon_{it} \tag{1}$$

Where Y_{it} corresponds to the outcomes I am going to evaluate; D_{i2015} is what is known as a policy variable, that is equal to one from 2015 onwards if a school, i, had a selected SPP student in the first wave of SPP, and zero otherwise; D_{i2015} is always zero for never treated schools. π_t are year fixed effects, π_i are school fixed effects and X_{it} is a vector of observed characteristics and ε_{it} are standard errors clustered at the level of the school. I use fixed effects by year and schools to control for systematic differences across schools, thus mitigating selection bias. I also control for some school characteristics such as women ratio, victims of the conflict ratio, average age, and average socioeconomic strata as a proxy to control for average income.

Even though socioeconomic strata is commonly used as a proxy for income or poverty, I construct an alternative proxy that identifies the average number of assets in student's families using variables reported by ICFES in the Saber 11 report about tv, computer, washing machine, microwave, and car ownership. I present these alternative results in Appendix V. arriving at the same conclusions.

The parameter of interest is β_1, which measures the change in the outcome variables of the schools with selected students compared to the ones that have not had a selected student, conditional on the set of school and year fixed effects.

[12] Results available upon request

As explained above, in my base definition of treatment, regarding the extensive margin, I define schools that had at least one SPP student in the first wave of SPP as treated. Hence, D^1_{i2015} takes the value of 1 for those schools for all years from 2015 onwards, the year after those schools had SPP students for the first time.

In addition to the base definition, I add two alternative definitions of the treatment variable that allow me to explore the effects in the intensive margin.D^2_{i2015} is the number of SPP selected students in 2014 and D^3_{i2015}, being the proportion of selected students compared to the total students of 11th grade in 2014. These variables are continuous and are positive from 2015 onwards, since the main assumption is that having SPP students in the previous year, 2014, potentially generated incentives that have an effect on schools achievement outcomes. D^j_{i2015} is always zero for never treated schools.

The first SPP students in 2014 were selected after taking the test and at the end of the academic year, preventing manipulation. The program was launched and announced by the government two months before the end of the school year in 2014 and after S11 were taken, preventing anticipatory effects on any of the outcomes in 2014. Then, the potential effects of the program should be seen from 2015 onwards.

Furthermore, the attribution of causality of my estimates relies on the assumption that there were not any other public programs or events in 2015 that affected the academic results of treated schools differently from never treated. As stated by Laajaj et al (2018) SPP is overwhelmingly perceived as the major change in education that occurred in Colombia and, as far as my knowledge, there are not any other public programs affecting schools with SPP students in 2014.

To attribute causality of the program to my estimates, I test the parallel trend assumption, that is, in the absence of selected students, the changes in educational outcomes before and after 2015 would have been similar among schools. For this purpose, I use a similar specification dividing the policy variable into interactions of the treatment variable with every year dummy during the analyzed period, presented in detail in the following section.

Another potential concern is the possibility that it exists spillover effects between treated schools and non-treated surrounding schools. Schools in the same area could exchange information, compete between them, and face some kind of peer effect facing similar incentives to the ones that are treated. Moreover, there can be networks of students and parents. Hence, for those students even if they are in a non-treated school incentives are going to be stronger that for never treated schools that are not exposed to treated schools, although one may think incentives are not as strong as for directly treated schools according to the discussion presented. In the specification presented in this section, those schools that are potentially facing a spillover effect are included in the control group, reducing the estimated effect. Therefore, while I cannot address directly the spillover concern my results could been understood as a lower bound.

After presenting the results for the extensive and intensive margin, I look for potential heterogeneous effects according to the proportion of need-based eligible students in each school. While I have no access to the exact SISBEN score of every student in each school, I have access to information about the number of students that meet the need-based criterion of SPP. Then, I use the proportion of students that have a SISBEN score under the program's cutoff value in the school; this variable interacts with D^1_{i2015}. I expect higher positive effects in the schools that have more potential SPP students. For these students, incentives should be stronger and schools could target a wider number of students to focus their efforts on test preparation trying to increase the number of SPP selected students.

6 Results

Regarding students, the treatment could have a positive, inspirational impact that leads them to choose higher effort levels and dropping out less. Thus, one would expect that results show, at the aggregate level of school, a decrease in dropout rates across grades. Also one would expect better academic results especially in S11 where incentives are stronger. The qualitative evidence found by prior evaluations of the program discussed in the context section suggest that the program generate bigger aspirations for parents and students in line with new incentives emerging. In addition the results found by Laajaj et al (2018) provide evidence than for 11 graders the incentives are stronger among need elegible high skilled population. However, the incentive effect to invest more effort in education is weaker for younger students in lower grades. For that reason, the effect of the program in academic results for S9 and S5 is not obvious. Moreover, if the program causes less dropouts, marginal students choosing to remain studying in the presence of the program are more likely to have low skills / low performance. In that sense, the effect on S5 and S9 could be negative if it is explained by a change in the composition of skills distribution in those grades now that low skill students chose not to drop out.

Regarding the incentives of schools, it is possible that schools improve their quality and teaching strategies, and this should be correlated with better academic results and lower dropout rates. On the other hand, there is a possibility that by trying to maximize the number of SPP students, schools are strategically focusing efforts on teaching specific S11 test-taking skills, as suggested by some interviews performed among institutions that provide tests preparation[13]. Moreover, those efforts could have increased specially for 11 grade students, need-based eligible students, and best performer high skill students. In this sense, the aggregated effect of the incentives generated by the program on school results is mixed since it could be benefiting some students and not benefiting others, and the net aggregated effect will depend on the distribution of these populations inside the school[14].

Nonetheless, the current data does not allow me to separate the effect of incentives of SPP on academic out-

[13] This exercise is discussed in detail in the context section. Audios available upon request

[14] Moreover, the net effect will depend on the marginal return of teacher and school attention, efforts, etc. in these populations

comes generated by incentives on students and by incentives on schools, nor does it allow me to completely identify each one of the channels described above. However, the data provide evidence of the combined effect of the incentives generated by SPP on students and schools regarding education achievement outcomes. Moreover, I present some evidence supporting some of the potential mechanisms in which the incentive scheme generated by SPP could have affected academic results.

6.1 Difference in differences results

Table 3. reports the results of the difference-in-differences analysis described in model (1) on standardized tests results. The coefficient in the first row corresponds to the estimation of β_1. Having an SPP student in the previous year has a positive and significant effect of a 0.058 standard deviation on the standardized math score of saber 11. However, the effect of the program on the standardized language score of S11 is not significant.

Regarding S9, the program had a negative and significant effect on both math and language standardized test scores. Schools with selected SPP students in the first year of the program have significantly lower results in Saber 5 in math and language scores than schools that never had SPP students.

As stated before, these results are robust when using a control group that only requires control schools to have at least one student above S11 cutoff value of 2014 in 2015 presented in Appendix IV. and when I replace the control of socioeconomic strata for the average assets variable described in the previous section presented in Appendix V. In addition, I find similar results when the control group is the totality of never treated schools, when I used a balanced panel for S9 and S5 test results[15] and when I perform a difference in differences with multiple treatment timing model presented in Appendix III.

Motivational positive effects seemed to be stronger for 11th grade students than for younger students. The results suggests the positive motivational effect is not strong enough for grades before the last year of high school to make the combined effect of the program on schools and students to be positive on aggregated academic results[16].

[15] Results avaliable upon request

[16] for instance, the motivational effect is very low for 9 and 5 graders or is affecting a reduced part of the population and the effect of recomposition of skills due to lower dropouts is bigger

Table 3: Difference in differences effect of SPP on standardized test results

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D^1_{i2015} - Dummy if school had SPP	0.0582***	0.0180	-0.118***	-0.0524***	-0.0544**	-0.0499**
the first year of the program	(0.0201)	(0.0200)	(0.0220)	(0.0201)	(0.0246)	(0.0211)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,567	15,567	15,567	15,567	15,567	15,567
R-squared	0.891	0.902	0.865	0.887	0.799	0.847

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Looking for the effect of SPP on the intensive margin following the treatments stated in Section 5, in Table 4., I present the results for the models using D^2_{i2015} and D^3_{i2015}. For conciseness, as in Table 3., I only report the estimated coefficient of interest β_1 and its standard error.

Considering the treatment defined as the number of SPP students, the results are consistent with the results obtained in the baseline treatment. Moreover, this contributes to the evidence that schools have a different response in the years after the launch of the program when they had selected students. An additional SPP student is significantly associated with higher scores of S11 Math score, lower S9 math scores and lower S5 test scores in both math and language. However, I find no robust results when using the proportion of 11th grade students selected for SPP as treatment, what can be explain by a very low variance in the proportion of SPP since is very low for most schools.

Table 4: Difference in differences effect of SPP on standardized test results – Intensive margin treatments

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D^2_{i2015} - Number of SPP	0.00523**	-0.000282	-0.00979***	-0.00234	-0.00508**	-0.00292*
the first year of the program	(0.00248)	(0.00156)	(0.00199)	(0.00154)	(0.00202)	(0.00169)
D^3_{i2015} - % of SPP	0.386	0.326*	-0.425	-0.584**	-0.122	0.0880
the first year of the program	(0.244)	(0.196)	(0.276)	(0.282)	(0.293)	(0.213)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,567	15,567	15,567	15,567	15,567	15,567
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regarding the effect of the program on dropout rate, Table 5. reports the results of the coefficients for the baseline treatment and the alternative measures of treatment for the aggregate dropout rate of the school and for each grade of high school. The effect of SPP on the dropout rate is negative and significant for all treatment definitions. The effect is also negative and significant for all lower secondary grades and for 10th grade, and not significant in 11th grade; that is not strange, since students who have already reached 11th grade have strong incentives to finish the academic year and 11th grade dropout rates are usually low.

Although without individual data I cannot completely verify the hypothesis that a change in composition is affecting negatively the aggregated test results for S5 and S9 by retaining low skill students, these results present some evidence leaning towards that hypothesis.

Table 5: Difference in differences effect of SPP on Dropout rate

	General	6th grade	7th grade	8th grade	9th grade	10th grade	11th grade
D^1_{i2015} - Dummy if school had SPP	-0.00568***	-0.00820***	-0.00480***	-0.00684***	-0.00329*	-0.00454**	-0.000120
the first year of the program	(0.00111)	(0.00200)	(0.00176)	(0.00172)	(0.00173)	(0.00225)	(0.00151)
D^2_{i2015} - Number of SPP	-0.000388***	-0.000484***	-0.000383**	-0.000496***	-0.000353**	-0.000381**	-0.0000492
the first year of the program	(0.000105)	(0.000160)	(0.000157)	(0.000161)	(0.000147)	(0.000191)	(8.84e-05)
D^3_{i2015} - Proportion of SPP	-0.0444***	-0.101***	-0.0670***	-0.0583***	-0.0429***	-0.0580***	-0.00264
the first year of the program	(0.00942)	(0.0161)	(0.0162)	(0.0175)	(0.0151)	(0.0169)	(0.0115)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,497	15,497	15,497	15,497	15,497	15,497	15,497
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.2 Identifying assumption of parallel trend

For my research design to be valid, schools with and without an SPP selected student should follow parallel trends in the years before SPP first year. However, since there are only three years available in the pretreatment period, it may be difficult to distinguish between statistical noise and deviations from the common trends by analyzing raw data. To take into account the possibility of different pre-trends for the treated and control group of schools, I estimate the following equation following the approach used in Autor (2003):

$$Y_{it} = \sum_{k=2012}^{k=2013} \beta_k D_{ik} + \sum_{k=2015}^{k=2017} \beta_k D_{ik} + \alpha X_{it} + \pi_t + \pi_i + \varepsilon_{it} \tag{2}$$

where Y_{it} is the outcome for individual i at time t, π_t and π_i are time and individual fixed effects, and X_{it}

is the same set of covariates as in equation (1). with D_{ik} I am including interactions of the time dummies and the treatment indicator for each year in my sample, omitting the pre treatment year 2014. That is, all the other interactions are expressed relative to the omitted period which serves as the baseline as is commonly used in literature. If the outcome trends for the treatment and control schools are the same in the period before SPP, then β_{2012} and β_{2013} should be non-significant. Moreover, even when significant, the inclusion of these variables is helping to control for pre-treatment differences in my estimates.

Table 6: Assessing parallel trends assumption for baseline treatment

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language	Dropout rate
D_{2012}	-0.0349	-0.00407	0.0390	0.0426	0.0352	0.0408	-0.0117**
	(0.0310)	(0.0296)	(0.0377)	(0.0354)	(0.0412)	(0.0373)	(0.00519)
D_{2013}	-0.0500*	-0.0578**	0.0106	-0.0127	-0.0373	-0.00516	-0.00351**
	(0.0269)	(0.0275)	(0.0321)	(0.0271)	(0.0356)	(0.0342)	(0.00159)
D_{2014}	-	-	-	-	-	-	-
D_{2015}	0.0228	-0.0385*	-0.0716**	-0.0245	-0.0568	-0.0352	-0.00581***
	(0.0251)	(0.0231)	(0.0343)	(0.0299)	(0.0392)	(0.0364)	(0.00206)
D_{2016}	-0.0553*	0.0261	-0.0592*	-0.0284	-0.000297	-0.00103	-0.00913***
	(0.0284)	(0.0248)	(0.0344)	(0.0294)	(0.0364)	(0.0309)	(0.00179)
D_{2017}	0.126***	-0.0191	-0.174***	-0.0820***	-0.127***	-0.0911***	-0.0101***
	(0.0314)	(0.0272)	(0.0327)	(0.0284)	(0.0357)	(0.0312)	(0.00176)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,567	15,567	15,567	15,567	15,567	15,567	15,567
R-squared	0.892	0.902	0.865	0.887	0.799	0.847	0.517

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In Table 6., I present these estimations for the baseline definition of treatment. Moreover, Figure A.III. in Appendix VI. shows the estimated coefficients of interaction graphically. As it could be seen, the parallel trend assumption holds for S5 and S9 in both math and language scores, the negative effect increases through the years after the beginning of the program. In the case of S11, the coefficients associated with policy interactions in the pre-treatment period are different from zero. However, for those years, schools that had SPP students in the treatment period have lower results than schools in the control group; then according to estimations presented in section 6.1 having an SPP program is reversing that trend and actually significantly changing the sign for math scores on average in the post treatment period. The assumption does not hold for the dropout rate. In Appendix VI., I also present the same estimates for the alternative definitions of treatment finding similar results[17].

Nevertheless, to provide further evidence on the conclusion regarding the effect of SPP on education achievement outcomes at the school level, I rebalance the sample to create groups that are more comparable in the pre-treatment period and present the results leading to the same conclussions in Appendix VII. To accomplish that I use a multivariate reweighting method presented in Hainmueller (2012) to reweight control group

[17] Similar results are also obtained in the alternative control groups mentioned in section 5

data to match the pre-treatment variable moments in the treatment group. In appendix VII. I also present the graphical results of the estimations when using the model described in equation (2). This strategy confirms the negative effect of having one SPP student for standardized tests results in both language and math in S5 and for math in the case of S9. I also find a positive effect of having a Selected SPP student on S11 and a negative effect on the dropout rate when finding parallel pre-trends in the rebalanced sample.

6.3 Heterogenous effects

While the effect of SPP has effects on the aggregate of schools' students, it is probable that it would affect some students more than others. For instance, need-based eligible students would have a stronger motivational effect since they already meet one of the conditions for the program. Incentives here are not only generated by the peer effect, the motivational effort to perform similarly as their peers, or to perform better in order to get different possibilities to go college, but through a more direct incentive that performing above the cutoff value of S11 will assure an SPP scholarship for them.

On the other hand, in the case when schools are devoting additional efforts for the need-based eligible population, schools with higher proportion of selected students will dedicate extra efforts to a bigger proportion of their total students related with better academic results at the aggregated school level.

Therefore, I look for potential heterogeneous effects according to the proportion of need-based eligible students in each school. Using aggregated data from the Ministry of Education for the years 2013 to 2016, I am able to identify how many students in the school have a SISBEN score that is below the defined cutoff value of the program; hence, I can calculate the proportion of eligible students at the school level.

Following the intuition proposed on model 1, I estimate the same regression adding an interaction term of the policy variable D1 with the proportion of eligible students. Therefore, I am able to calculate the marginal effect of having at least one SPP selected student on the first year of the program for different values of the proportion of need based eligible students.

$$Y_{it} = \beta_1 D_{i2015} + \beta_2 need\_elegible_{it} + \beta_3 D_{i2015} * need\_elegible_{it} + \alpha X_{it} + \pi_t + \pi_i + \varepsilon_{it} \qquad (3)$$

In Appendix VIII., I report the results of this regression. However, for a visual interpretation in Figure 3 - 6, I present the marginal effect of having an SPP student according to the proportion of eligible students for tests results and dropout rate.

Figure 2: Heterogeneous effects of D^1_{i2015} according to the proportion of need-need based eligible students - **Saber 11**

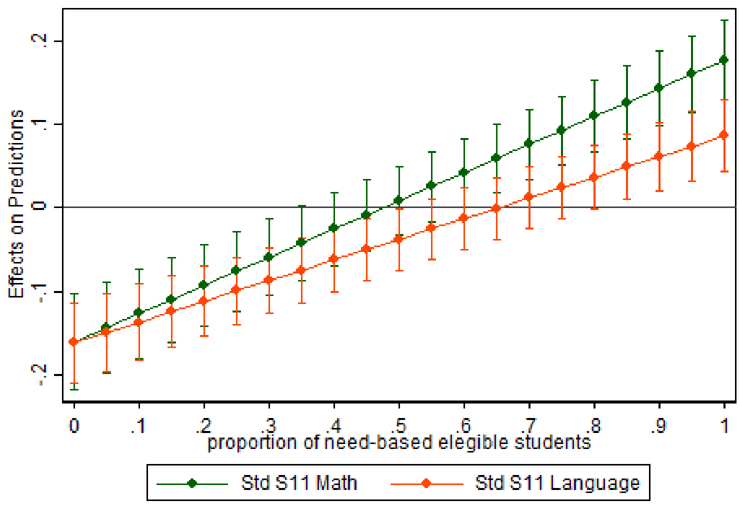


Figure 3: Heterogeneous effects of D^1_{i2015} according to the proportion of need-need based eligible students - **Saber 9**

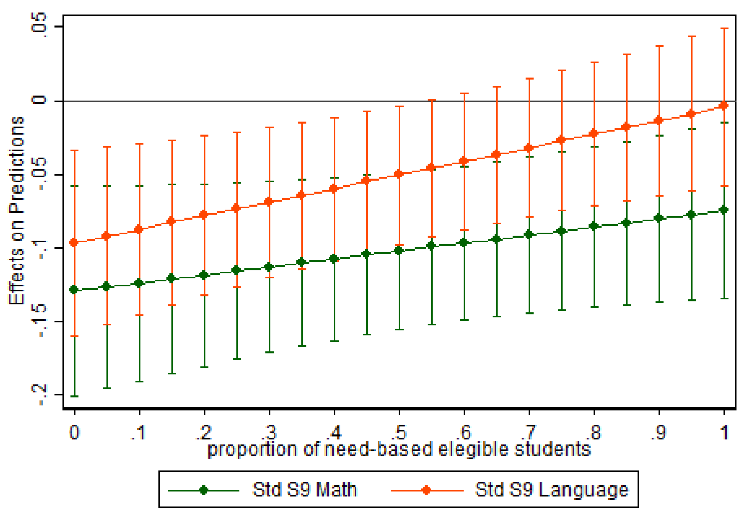


Figure 4: Heterogeneous effects of D^1_{i2015} according to the proportion of need-need based eligible students - **Saber 5**

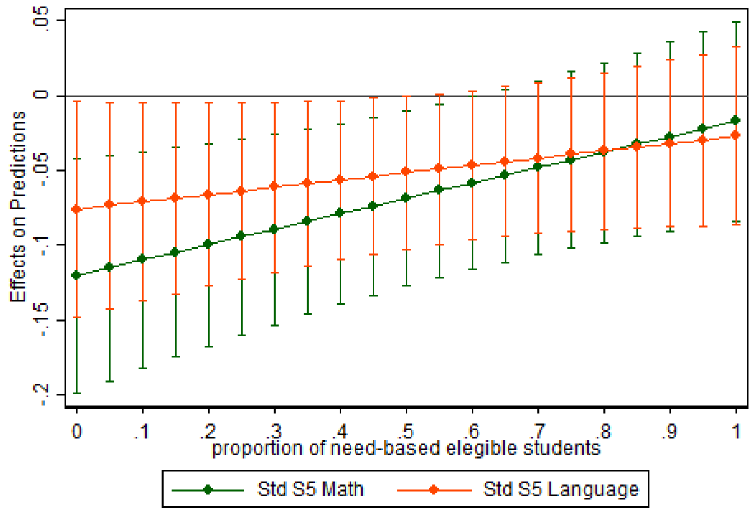
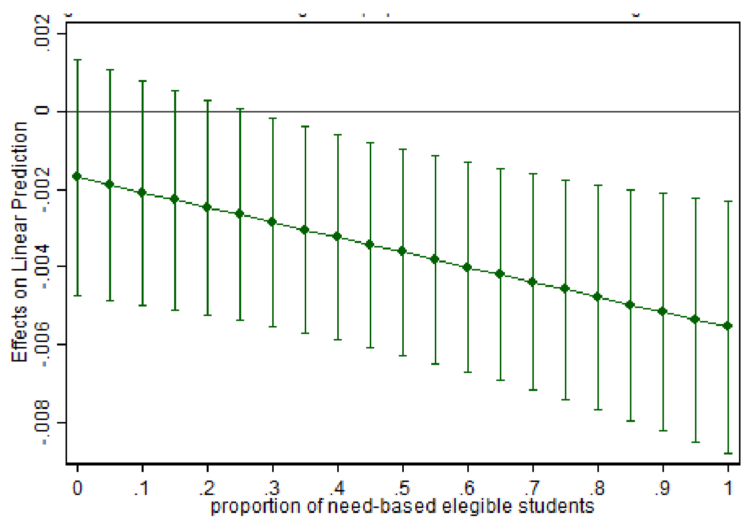


Figure 5: Heterogeneous effects of D^1_{i2015} according to the proportion of need-need based eligible students - **Dropout rate**



As expected, for S11 test results, the marginal effect of having an SPP student for schools with low proportion of need-based eligible students is lower and even negative. Instead, for schools with high proportions of need-based eligible students this effect is positive. The rise in the marginal effect of the treatment is higher for standardized math score in comparison with the language score. This supports the hypothesis that SPP is affecting differently aggregated schools depending on the composition of the elegible population. The positive effect is only focused in schools with high proportion of need eleigible students for saber 11 tests scores. Moreover, The effect for schools with low proportion of need-eligible students is negative.

I find similar results for S9 and S5 with higher proportions of need-eligible students reducing the negative effect of having a selected SPP student. However in this case, the marginal effect of D^1_{i2015} is not significantly different between schools with low and high proportion of need-based eligible students. I also find that having a higher proportion of SPP students is related with a stronger negative effect on the dropout rate but the effect is not significantly different for schools between schools with high or low proportion. This supports the hypothesis that for younger students the concentration of motivational effect among different populations is not as important as the effect of the potential change in the skills composition because students are changing dropout decision.

7 Discussion

At the end of 2014 the government of Colombia launched the public program SPP a need-and-merit-based scholarship program for going to college. The program provides a good example of unexpected incentive schemes that arise when public policies affect larger groups than the targeted beneficiaries. The program created incentives in schools, parents and mainly in students that affect educational achievement outcomes. My analysis attempts to identify the causal effect of the SPP program on education achievement outcomes at the aggregated school level.

I find that incentives were stronger for schools with SPP students in 2014 in comparison with schools that had never had selected students. SPP has a positive robust effect on Saber 11 math scores. Having one SPP student has a negative effect in Saber 9 and Saber 5 tests scores, showing that gains in performance at the school level are concentrated only in the last year of school. SPP is related with lower dropout rates at the school level and across high school grades. Moreover, all the effects described above grow with the school's proportion of need-based eligible students.

While it is true that SPP appears to improve student achievement in 11th grade, it also has some potential drawbacks. The negative effect on Saber 9 and Saber 5 are partially explained by economic theory in the presence of short run discounting for younger students. It could also be explained by a change in the composition of skills as suggested by a decline in dropouts. Although I could not explore the potential mechanisms in detail in this document, it is important to continue disentangling the incentive schemes caused by these kinds of programs to improve the design of public policies. With a deeper understanding of these mechanisms, policymakers will be able to adapt this program or similar ones to enhance its benefits and mitigate its costs.

References

Albert, D., Chein, J., Steinberg, L. (2013). The teenage brain: Peer influences on adolescent decision making. Current directions in psychological science, 22(2), 114-120.

Álvarez, María José., Javier Corredor, Carlos Castro., Londoño, Juliana., Maldonado, Carolina., Rodríguez, Catherine., Tatiana Velasco, Fabio Sánchez., Ángel, Daniel Mateo., Ayala, María Camila., Pulido, Xiomara. (2017). El Programa Ser Pilo Paga: Impactos Iniciales en Equidad en el Acceso a la Educación Superior y el Desempeño Académico. Documentos CEDE

Angier N. (2009). A Molecule of Motivation, Dopamine Excels at Its Task. The New York Times. 2009 October 26; B1 and ff.

Angrist, J. D., Lavy, V. (2002). The effect of high school matriculation awards: Evidence from randomized trials (No. w9389). National Bureau of Economic Research.

Angrist, J., Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. American economic review, 99(4), 1384-1414.

Angrist, J., Lang, D., Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. American Economic Journal: Applied Economics, 1(1), 136-63.

Autor, David, (2003), Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing, Journal of Labor Economics, 21, issue 1, p. 1-42.

Baez, J. E., Camacho, A. (2011). Assessing the long-term effects of conditional cash transfers on human capital: evidence from Colombia. The World Bank.

Baker, G. P. (1992). Incentive contracts and performance measurement. Journal of political Economy, 100(3), 598-614.

Banerjee, A. V., Cole, S., Duflo, E., Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. The Quarterly Journal of Economics, 122(3), 1235-1264.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S. Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. Journal of Economic Perspectives, 31(4), 73-102.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., Perez-Calle, F. (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia. American Economic Journal: Applied Economics, 3(2), 167-95.

Behrman, J. (2010), Investment in Education Inputs and Incentives in Jere Behrman (eds.), Handbook of Development Economics, vol 5, Elsevier.

Belloni, A., Chernozhukov, V., Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. Journal of Economic Perspectives, 28(2), 29-50.

Belot, Michele, Erik Canton, and Dinand Webbink. (2007). Does reducing student support affect scholastic performance? Evidence from a Dutch reform. Empirical Economics 32 (2-3):261–275.

Bertrand, M., E. Duflo, and S. Mullainathan. (2004). “How much should we trust differences-indifferences estimates?” The Quarterly Journal of Economics, 119, 249–275. 4.1.1

Bettinger, E. (2004). How financial aid affects persistence. In College choices: The economics of where to go, when to go, and how to pay for it (pp. 207-238). University of Chicago Press.

Bettinger, E. (2005). Private School Vouchers in Colombia. In Draft paper prepared for the conference, Mobilizing the Private Sector for Public Education, World Bank and Kennedy School of Government, Harvard University, Cambridge, MA.

Bettinger, Eric P. (2012). "Paying to learn: The effect of financial incentives on elementary school test scores." Review of Economics and Statistics 94.3: 686-698.

Binder, M., Ganderton, P. T., Hutchens, K. (2002). Incentive effects of New Mexico’s merit-based state scholarship program: Who responds and how?. Who should we help, 41-56.

Borusyak, Kirill, and Xavier Jaravel. (2017). "Revisiting event study designs." Available at SSRN 2826228.

Brickman, P., Campbell, D., (1971). Hedonic relativism and planning the good society. In: Appley, M.H. (Ed.), Adaptation-level Theory: A Symposium. Academic Press, New York, pp. 287–302.

Bugler, D. T., Henry, G. T., Rubenstein, R. (1999). An Evaluation of Georgia’s HOPE Scholarship Program: Effects of HOPE on Grade Inflation. Academic Performance and College Enrollment, Council for School Performance, Georgia State University, Atlanta, GA.

Camacho, A. Messina, J. and Uribe, J. (2016). the Expansion of Higher Education in Colombia: Bad Students or Bad Programs? IDB discussion paper Nº IDB-DP-452.

Casey BJ, Jones RM, Hare TA Ann. (2008). Adolescent brain development: a period of vulnerabilities and opportunities. N Y Acad Sci

Centro Nacional de Consultoria and Centro de Estudios Sobre Desarrollo Económico CEDE - Universidad de los Andes. (2017). Informe de resultado del levantamiento de línea de base, evaluación de impacto de corto plazo y tercera entrega de la documentación de las bases de datos.

Colonnelli, E., Prem, M. (2017). Corruption and firms: Evidence from randomized audits in brazil.

Cohodes, S.R. and J.S. Goodman, (2014). “Merit Aid, College Quality, and College Completion: Massachusetts’ Adams Scholarship as an In-Kind Subsidy,” American Economic Journal: Applied Economics, 2014, 6 (4), 251–285.

Cornwell, C., Mustard, D. B., Sridhar, D. J. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia’s HOPE program. Journal of Labor Economics, 24(4), 761-786.

Cornwell, Christopher, Kyung Hee Lee, and David B. Mustard. (2003). The Effects of Merit-Based Financial Aid on Course Enrollment, Withdrawal and Completion in College. In IZA Discussion paper

Cullen, Julie Berry and Randall Reback,(2002) “Tinkering Toward Accolades: School Gaming under a Performance Accountability System,” Working paper, University of Michigan.

Dahl RE, Ann. (2004). Triadic model of the neurobiology of motivated behavior in adolescence. N Y Acad Sci

De Paola, Maria, Vincenzo Scoppa, and Rosanna Nisticò. (2012) "Monetary incentives and student achievement in a depressed labor market: results from a randomized experiment." Journal of Human Capital 6.1: 56-85.

Deci, E. L. (1972a). The effects of contingent and noncontingent rewards and controls on intrinsic motivation. Organizational behavior and human performance, 8(2), 217-229.

Deci, E. L. (1972b). Intrinsic motivation, extrinsic reinforcement, and inequity. Journal of personality and social psychology, 22(1), 113.

Deming, D., Dynarski, S. (2010). College aid. In Targeting investments in children: Fighting poverty when resources are limited (pp. 283-302). University of Chicago Press.

Departamento Nacional de Planeación. (2015). Todos por un Nuevo País.

Duarte, Jesús., Bos, María Soledad., Moreno, José Martín. (2012). Quality, Equality and Equity in Colombian Education (Analysis of the SABER 2009 Test). Inter-American Development Bank

Ernst, M., Pine, D. S., Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. Psychological medicine, 36(3), 299-312.

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

Fetler, M. (1989). School dropout rate, academic performance, size, and poverty: Correlates of educational reform. Educational Evaluation and Policy Analysis, 11, 109–116.

Figlio, D., Getzler, L. (2002). Accountability, ability and disability: Gaming the system?. National Bureau for Economic Research Working Paper 9307.

Frederick, S., Loewenstein, G., (1999). Hedonic adaptation. In: Kahneman, D., et al. (Eds.), Well-being: The Foundations of Hedonic Psychology. Russell Sage Foundation, New York, pp. 302–329.

Fryer Jr, Roland G. (2011). "Financial incentives and student achievement: Evidence from randomized trials." The Quarterly Journal of Economics 126.4: 1755-1798.

Glaeser, Edward L., and Andrei Shleifer. (2001) "Not-for-profit entrepreneurs." Journal of public economics 81.1 : 99-115.

Glewwe, P., Ilias, N., Kremer, M. (2003). Teacher Incentives" NBER Working Paper: No. 9671.

Golsteyn, B. H., Grönqvist, H., Lindahl, L. (2014). Adolescent time preferences predict lifetime outcomes. The Economic Journal, 124(580), F739-F761.

González-Velosa, C., Rucci, G., Sarzosa, M., Urzúa, S.(2015). Returns to higher education in chile and colombia [IDB Working Paper Series No. IDB-WP-587]. InterAmerican Development Bank.

Gregory J. Koop, Joseph G. Johnson. (2010). The use of multiple reference points in risky decision making. Behavioral Decision Making journal. Volume 25, Issue 1 Pages 49–62

Grissmer, D.W. et. al. (2000). Improving Student Achievement: What NAEP Test Scores Tell Us. MR-924-EDU. Santa Monica: RAND Corporation.

Guarín, A., Londoño, S., Medina-Durango, C. A., Parra, J., Posso-Suárez, C. M., Vélez, C. E. (2016). Estimating the effect of attending a public versus a private university in Colombia on academic achievement. Borradores de Economía; No. 968.

Hainmueller J (2012). “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies.” Political Analysis, 20(1), 25–46.

Hanushek, Eric A., and Margaret E. Raymond. (2005) "Does school accountability lead to improved student performance?." Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management24.2 : 297-327.

Holmstrom, B., Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. JL Econ. Org., 7, 24.

ICFES - Instituto Colombiano para la Evaluación de la Educación. 2011. SABER 5º y 9º (2009). Informe técnico de resultados históricos. Bogotá, D.C.

Jacob, B Levitt, S. (2003). "Catching Cheating Teachers: The Results of an Unusual Experiment in Implementing Theory," Brookings-Wharton Papers on Urban Affairs, vol 2003(1), pages 185-220.

Jacob, Brian A. (2002). "The Impact of Test-Based Accountability in Schools: Evidence from Chicago." John F. Kennedy School of Government, Harvard University.

Jacob, Brian A. and Steven D. Levitt. (2003) "Rotten Apples: An Investigation Of The Prevalence And Predictors Of Teacher Cheating," Quarterly Journal of Economics, v118(3,Aug), 843-878.

Kahneman, D. and Tversky, A. (1979). 'Prospect theory: an analysis of decisions under risk', Econometrica, vol. 47(2), pp. 263–91.

Kahneman, D., (1999). Objective happiness. In: Kahneman, D., et al. (Eds.), Well-being: The Foundations of Hedonic Psychology. Russell Sage Foundation, New York, pp. 3–25.

Kaldor, Nicholas. (1957). "A model of economic growth." The economic journal 67.268 (1957): 591-624.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. (2015). "Self-control at work." Journal of Political Economy 123.6: 1227-1277.

Klein, S. P., L. S. Hamilton, et al. (2000). "What Do Test Scores in Texas Tell Us?" Santa Monica, CA: RAND.

Kremer, Michael. (2003) "Randomized evaluations of educational programs in developing countries: Some lessons." American Economic Review 93.2: 102-106.

Kruglanski, A. W., Friedman, I., Zeevi, G. (1971). The effects of extrinsic incentive on some qualitative aspects of task performance 1. Journal of Personality, 39(4), 606-617.

Laajaj, R., Moya, A., Sánchez, F. (2018). Equality of Opportunity and Human Capital Accumulation: Motivational Effect of a Nationwide Scholarship in Colombia (No. 016352). UNIVERSIDAD DE LOS ANDES-CEDE.

Lan, W. Lanthier, R. (2003) Changes in Students' Academic Performance and Perceptions of School and Self Before Dropping Out of Schools, Journal of Education for Students Placed at Risk (JESPAR), 8:3, 309-332, DOI: 10.1207/S15327671ESPR0803$_2$

Lazear, E. P. (2000). The power of incentives. American Economic Review, 90(2), 410-414.

Leuven, Edwin, Hessel Oosterbeek, and Bas van der Klaauw. (2003). The Effect of Financial Rewards on Students' Achievements: Evidence from a Randomized Experiment. In CEPR Discussion paper

Leuven, Edwin, Hessel Oosterbeek, and Bas Van der Klaauw. (2010). "The effect of financial rewards on students' achievement: Evidence from a randomized experiment." Journal of the European Economic Association 8.6: 1243-1265.

Levitt, Steven D., List, John A. Neckermann, Susanne. Sadoff,. Sally. (2016) "The behavioralist goes to school: Leveraging behavioral economics to improve educational performance." American Economic Journal: Economic Policy 8.4: 183-219.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. Journal of Labor Economics, 28(4), 825-860.

Locke, Edwin A., and Gary P. Latham. (1990) A theory of goal setting task performance. Prentice-Hall.

Londono-Velez, J., Rodriguez, C., Sánchez, F. (2018). Upstream and Downstream Impacts of College Merit-Based Financial Aid for Low-Income Students: Ser Pilo Paga in Colombia. American Economic Journal: Economic Policy.

Loomes, Graham, and Robert Sugden. (1987). "Testing for Regret and Disappointment in Choice under Uncertainty." Economic Journal, 97(S): 118–29.

MacLeod, W.B. Riehl, E., Saavedra, J.E., and M. Urquiola. (2017). "The Big Sort: College reputation and labor market outcomes," American Economic Journal: Applied.

McPherson, M. S., Schapiro, M. O. (2006). College Access: Opportunity or Privilege?.

Mealli, F., Rampichini, C. (2012). Evaluating the effects of university grants by using regression discontinuity designs. Journal of the Royal Statistical Society: Series A (Statistics in Society), 175(3), 775-798.

Ministerio de Educación Nacional. 2018. Informe de gestión 2014 – 2018.

Ministerio de Educación. (2006). Cartilla 1: el proceso de matrícula.

OECD (2016), Education in Colombia, Reviews of National Policies for Education, OECD Publishing, Paris.

OECD and The World Bank, (2012). Reviews of National Policies for Education: Tertiary Education in Colombia, OECD Publishing, 2012.

Okun, M. A. (1976). Adult age and cautiousness in decision. Human Development, 19(4), 220-233.

Riehl, E., J.E. Saavedra, and M. Urquiola, 2016. Learning and earning: An approximation to college value added in two dimensions," October 2016. NBER Working Paper No. 22725

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. American Educational Research Journal, 32, 583–625

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

Schreck, C. Fisher, B. Miller, M (2004). The social context of violent victimization: A study of the delinquent peer effect, Justice Quarterly, 21:1, 23-47, DOI: 10.1080/07418820400095731

Scott-Clayton, J. (2011). On money and motivation a quasi-experimental analysis of financial incentives for college achievement. Journal of Human resources, 46(3), 614-646.

Scott-Clayton, J. and B. Zafar, (2016). “Financial Aid, Debt Management, and Socioeconomic Outcomes: Post-College Effects of Merit-Based Aid,” August 2016. NBER Working Paper 22574.

Sokatch, A. (2006). Peer influences on the college-going decisions of low socioeconomic status urban youth. Education and Urban Society, 39(1), 128-146.

Solis, A. (2017). Credit access and college enrollment. Journal of Political Economy, 125(2), 562-622.

Spear L. (2009). The behavioral neuroscience of adolescence.

Stiglitz, J. E. (1969). Distribution of income and wealth among individuals. Econometrica: Journal of the Econometric Society, 382-397.

Tokuno, K., Campbell, F. (1992). The freshman interest group program at the University of Washington: Effects on retention and scholarship. Journal of The First-Year Experience Students in Transition, 4(1), 7-22.

Tversky, A. and Kahneman, D. (1991). ‘Loss aversion in riskless choice: a reference-dependent model’, Quarterly Journal of Economics, vol. 106(4), pp. 1039–61.

Ulanowicz, M.A; Parra, K; Wendle, E; Tisiana, L. (2006). Riesgos en el embarazo adolescente. Revista de Posgrado de la Via Cátedra de Medicina, No 153.

Velasco, T. and Sanchez, F. (2014). Do Loans for Higher Education Lead to Better Salaries? Evidence from a Regression Discontinuity Approach for Colombia. ” Serie Documentos Cede, 2013-41 ISSN 1657-7191 Edición electrónica.

Velasco,T. Sanchez, F. (2014). Do loans for higher education lead to better salaries? Evidence from a regression discontinuity approach for Colombia," October 2014.

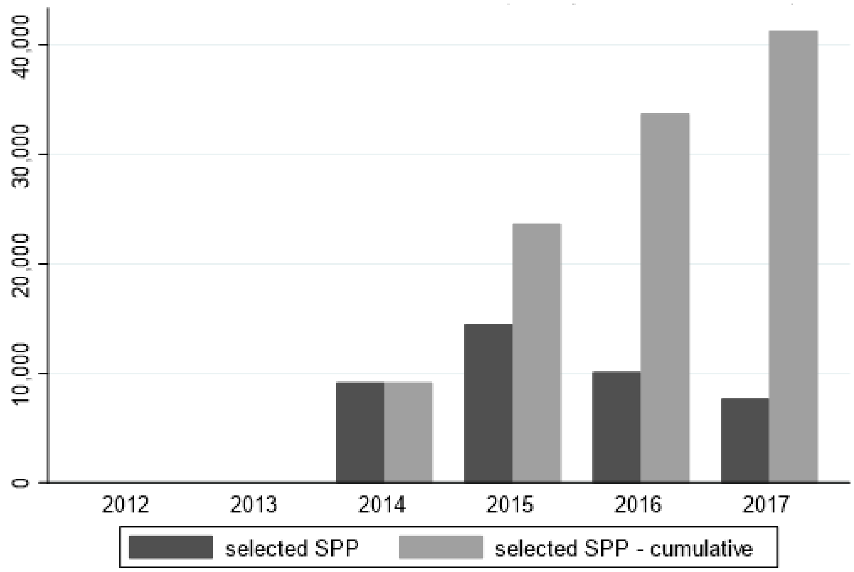
Wooldridge, Jeffrey M. (2005) "Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models." Review of Economics and Statistics 87.2 : 385-390.

Appendix

Table A.I. Descriptive statistics for outcome variables

Variable	Obs	Mean	Std. Dev.	Min	Max
		2012			
S11 Math	5,520.00	44.08	3.72	30.00	70.68
S11 Language	5,520.00	45.09	2.76	29.00	60.08
S9 Math	5,520.00	286.54	42.02	120.00	486.00
S9 Language	5,520.00	291.42	41.49	125.00	453.00
S5 Math	5,520.00	286.66	41.58	145.00	492.00
S5 Language	5,520.00	293.35	42.69	158.00	487.00
		2013			
S11 Math	6,641.00	44.71	5.46	22.00	82.23
S11 Language	6,641.00	47.02	3.94	32.00	71.00
S9 Math	6,641.00	298.93	53.53	170.00	500.00
S9 Language	6,641.00	297.58	53.11	156.00	500.00
S5 Math	6,641.00	304.63	49.91	170.00	500.00
S5 Language	6,641.00	308.90	51.57	167.00	500.00
		2014			
S11 Math	7,098.00	49.94	5.90	33.00	81.14
S11 Language	7,098.00	49.69	5.98	31.00	76.56
S9 Math	7,098.00	296.32	56.13	161.00	500.00
S9 Language	7,098.00	296.22	54.42	151.00	500.00
S5 Math	7,098.00	297.15	47.94	150.00	500.00
S5 Language	7,098.00	302.76	49.94	148.00	500.00
		2015			
S11 Math	7,854.00	50.55	7.94	27.25	92.40
S11 Language	7,854.00	49.92	5.83	33.33	76.24
S9 Math	7,854.00	299.46	58.35	159.00	500.00
S9 Language	7,854.00	296.55	58.92	150.00	500.00
S5 Math	7,854.00	307.69	51.32	100.00	500.00
S5 Language	7,854.00	304.27	51.28	169.00	500.00
		2016			
S11 Math	8,051.00	51.37	7.98	28.00	93.77
S11 Language	8,051.00	52.79	5.84	32.00	73.49
S9 Math	8,051.00	316.20	46.19	181.00	500.00
S9 Language	8,051.00	308.92	42.90	179.00	482.00
S5 Math	8,051.00	310.26	39.85	179.00	484.00
S5 Language	8,051.00	318.61	40.82	197.00	475.00
		2017			
S11 Math	9,007.00	50.01	8.29	1.77	81.07
S11 Language	9,007.00	53.12	6.32	8.00	86.00
S9 Math	9,007.00	310.62	45.75	153.00	500.00
S9 Language	9,007.00	313.96	39.76	194.00	480.00
S5 Math	9,007.00	304.30	40.92	187.00	500.00
S5 Language	9,007.00	316.68	38.56	214.00	500.00

Figure A.I. Number of students selected to SPP (each year and cumulative)



Appendix I. Collection of some news on the SPP in schools

“Teníamos mucha ilusión de ser los nuevos ‘Pilos’ del colegio”

El Distrito se mantuvo durante el último año entre las primeras ciudades a nivel nacional con más ‘Pilos’. En 2017 obtuvo 410 estudiantes beneficiarios con el programa, mientras que el municipio de Soledad contó con un total de 176 estudiantes acreedores de la beca. http://tiny.cc/wgc06y

Colegios de Barranquilla consiguen 632 becas ’Ser Pilo Paga’

De las 150 instituciones educativas distritales 85 tuvieron estudiantes beneficiados con el programa ’Ser Pilo Paga’ del Gobierno Nacional por su excelente desempeño en las Pruebas Saber, lo que es muestra de importantes logros en materia educativa pública que ha tenido Barranquilla en el último cuatrienio… rectora de la IED Madre Marcelina, destacó: "Este logro también es de todos los actores de la educación, quienes se comprometen a dar lo mejor de cada uno para que estos jóvenes hoy tengan oportunidades tan valiosas y se sigan formando con calidad"… De acuerdo con el Secretario de Educación Dagoberto Barraza, este logro es fruto del compromiso y el trabajo en equipo de docentes, rectores y padres de familia. http://tiny.cc/1ic06y

24 Colegios De Bolívar Recibieron Incentivos Económicos Por Más De 2 Mil 500 Millones De Pesos

El líder de Calidad Educativa de la Secretaría de Educación de Bolívar, Pedro Pulido Franco, aseguró que Bolívar sí avanzó en el mejoramiento de la calidad en los grados de primaria y secundaria, “vamos a trabajar conjuntamente con las políticas y programas que adelanta el Ministerio como Ser Pilo Paga, Todos Aprender, Plan de Formación Docente e iniciativas que adelanta el Gobernador http://tiny.cc/ak616y

¿Se acaban incentivos para colegios con mejores resultados en la Prueba Saber?

Durante dos años consecutivos, 59.787 docentes, directivos y administrativos de 2.354 establecimientos educativos en el país recibieron estímulos por sus resultados de desempeño de 2015 y 2016. El Gobierno Nacional, en cabeza del Ministerio de Educación (MEN), entregó 148.437 millones de pesos como incentivos a la calidad http://tiny.cc/3r616y

Evardo Turizo, el colegio que forma periodistas del conocimiento

Ha sido de esta manera como hemos logrado que entre 2014 y 2015, 26 de nuestros bachilleres fueran becados por el programa Ser Pilo Paga. Esperamos que por lo menos 30 de nuestros estudiantes sean premiados con este beneficio del Gobierno Nacional en el 2016. http://tiny.cc/st616y

Sociedad Pilo promoverá el Gen Ciudadano en 16 colegios del Atlántico

La Ruta de Activación del Gen Ciudadano consiste en un programa de mentoría, en el que universitarios beneficiarios del Programa Ser Pilo Paga, pertenecientes al grupo estudiantil Sociedad Pilo, se acercarán a los colegios seleccionados en Barranquilla y municipios de Atlántico para desarrollar laboratorios de formación ciudadana en los que participen los bachilleres. http://tiny.cc/bv616y

El cara y sello del programa "Ser Pilo Paga"-Los colegios destacados en Popayán

De otra parte, en Popayán 40 colegios lograron cupos en "Ser Pilo Paga", de estos San Agustín y Don Bosco alcanzaron 11 y 12 cupos respectivamente, los demás entre seis y un cupo. Entrevistamos a los rectores de estas instituciones para conocer cómo forman a sus estudiantes. Comenzaron a hacer simulacros de Pruebas Saber desde el grado 9, con la asesoría de una empresa caleña, recursos propios del colegio y aportes mínimos de padres de familia. Los resultados se convirtieron en indicadores de cómo están las estudiantes en el proceso educativo revelando fortalezas y falencias para luego diseñar planes de mejoramiento, estas son las estrategias pedagógicas para elevar el nivel. http://tiny.cc/bv616y

La fórmula del mejor colegio público de Colombia

Más de 30 docentes y 22 establecimientos educativos de varios departamentos fueron galardonados por el presidente Juan Manuel Santos y la ministra de Educación, Gina Parody, la noche del lunes, con incentivos económicos y tabletas como un homenaje a su labor. Varias categorías fueron premiadas, entre ellas el mejor maestro, la mejor institución pública y privada, la mejor escuela normal superior y la institución con más "pilos". De los resultados más recientes en las pruebas Saber, 42 alumnos de 60 tenían derecho a las becas para el programa Ser Pilo Paga del Gobierno, http://tiny.cc/7y616y

Estos son los profesores y colegios premiados por el Presidente

Durante la ceremonia, el Presidente y la Ministra exaltaron las prácticas pedagógicas innovadoras de las regiones que han fortalecido los aprendizajes de los niños y jóvenes. Además, hicieron un reconocimiento a las instituciones que han aprovechado de forma especial programas como la Jornada Única, el programa Todos a Aprender, el Índice

Sintético de Calidad Educativa, Ser Pilo Paga, el Día E y Bilingüismo. La quinta subcategoría es a la Mejor institución educativa con más 'Pilos'. Reconoce a la institución educativa del país que logró el mayor número de estudiantes potenciales beneficiarios de la tercera edición del programa 'Ser pilo paga': Institución Educativa Nuestra Señora del Rosario, de Colón (Nariño); Institución Educativa Leopoldo López Álvarez, de Génova (Nariño); y la Institución Técnica Nacional de Comercio, de Cúcuta (Norte de Santander). http://tiny.cc/o0616y

Gobierno premió a los mejores en educación durante 2017 Distinción a entidades territoriales La tercera distinción premió el trabajo de las entidades territoriales que promueven tanto el acceso como la calidad de la educación superior en sus regiones, y que invierten en proyectos de fomento como los fondos 'Pilo Regional'. Hoy fueron reconocidas las gobernaciones del Atlántico, San Andrés, Providencia y Santa Catalina y Valle del Cauca, por sus proyectos en formación técnica y tecnológica que, con el acompañamiento del Ministerio de Educación, están transformando la vida de cientos de jóvenes e impulsando la productividad de sus departamentos. http://tiny.cc/o1616y

Colegio de Medellín, el segundo en ganar becas Ser Pilo Paga

Hoy la institución saca pecho por ser la segunda del país y la primera en Antioquia con más estudiantes ganadores de becas Ser Pilo Paga. puso 178 estudiantes en este programa nacional, después de la Institución Educativa Ciudad de Pasto, donde 212 bachilleres ganaron el derecho a realizar estudios universitarios de manera gratuita. En 2017, del Inem ganaron el concurso 38 estudiantes, lo que significa que la IE casi que quintuplicó los beneficiarios, lo que para el rector, Fernando Carvajal Oquendo, es reflejo del modelo educativo institucional, que es diferente al resto de colegios del país. http://tiny.cc/j3616y

Un colegio de Montería, entre los más pilos del país

Pese a que el 68% de los potenciales beneficiarios de la tercera versión programa del Gobierno, Ser pilo paga, son de colegios públicos, el colegio de Montería que se ubicó entre los diez que más "Pilos" aporta a Colombia es una institución educativa privada que con orgullo hace parte del 32% restante: La Salle Montería. http://tiny.cc/r4616y

Ser pilo paga 2 en cifras - Bogotá Estos son los colegios de Bogotá con el mayor número de pilos: http://tiny.cc/96616y

Appendix II.

Table A.II. Distribution of schools between treated and control

Year		2012	2013	2014	2015	2016	2017	Total
Control	Absolut	219	402	385	700	700	664	3,070
	% schools	11.2	15.7	13.9	22.6	22.6	21.7	18.6
Treated	Absolut	1,743	2,166	2,393	2,393	2,393	2,393	13,481
	% schools	88.8	84.4	86.1	77.4	77.4	78.3	81.5
Total		6,416	8,860	8,982	9,208	9,265	9,472	52,203

Figure A.II. Distribution of the number of SPP selected students

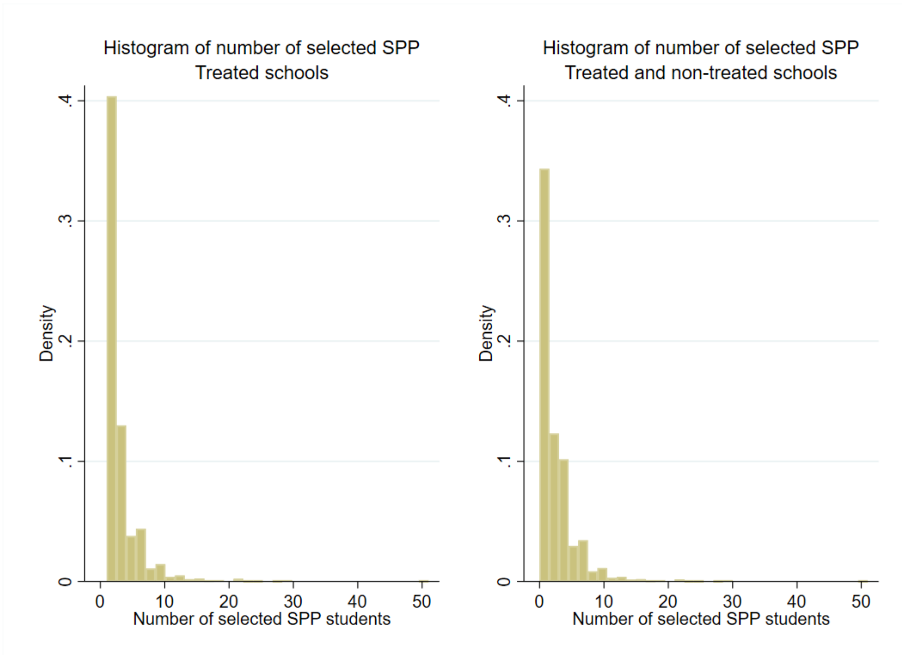


Table A.III. Descriptive statistics, Before 2015

Var	MeanT	MeanC	Diff	t	P-Value	ObsT	ObsC
Assets	2.73	2.88	-0.14	-3.50	0.00	6302	1006
Age	11.35	11.45	-0.10	-4.03	0.00	6302	1006
Rural	0.09	0.26	-0.17	-12.10	0.00	6302	1006
Public	0.80	0.57	0.22	13.60	0.00	6302	1006
Socioeco. strata	1.88	2.09	-0.21	-6.08	0.00	6302	1006
Women	0.51	0.49	0.03	6.87	0.00	6302	1006
Victim	0.03	0.03	0.00	-0.55	0.58	6302	1006
Prop. Of need eligible	0.49	0.41	0.08	6.94	0.00	4108	610
Size	1429.14	664.13	765.01	40.87	0.00	6302	1006
S11_Math	49.05	48.58	0.47	1.93	0.05	6302	1006
S11_Language	49.94	49.62	0.32	1.63	0.10	6302	1006
S9_Math	317.72	314.34	3.38	1.70	0.09	6302	1006
S9_Language	320.48	316.37	4.10	2.22	0.03	6302	1006
S5_Math	313.78	314.05	-0.27	-0.16	0.88	6302	1006
S5_Language	321.96	322.79	-0.84	-0.45	0.65	6302	1006
Dropout rate	0.03	0.03	0.00	-0.71	0.48	6302	1006

Appendix III. Difference in differences with multiple treatment timing

In this section, I estimating a dinamyc diff in diff redefining the treatment to start the year after schools received their first SPP student.

As mentioned in the empirical design, the approach of my main strategy is that for schools that had a SPP student at the end of 2014, it emerged an incentive scheme from 2015 onwards that could change the behavior of students and school staff and boards affecting the aggregated school academic results. In the main strategy, I compare those schools with never treated schools, that is to say, schools that never had a SPP student.

As explained in the section of context of SPP, it is possible that some schools had their first selected SPP student or students not in the first year of the program but in 2015 or 2016 or 2017, and hence the effects for those schools emerged at the next year. Those schools are excluded for the sample and represent the 25% of the observations. This exclusion is made to minimize the potential endogeneity between having an SPP in 2014 and in the following years[18].

In that sense, SPP framework could be interpreted as treatment that start in different years depending on the school. A school began to be treated the year after they had their first SPP student when I would expect to see stronger incentives and changes in behavior. In Table A.IX I present the percentage of schools that start to be treated for each year.

Table A.IV. percentage of schools that start to be treated each year

	% of schools
Never Treated	39.19
Treated from 2015	35.76
Treated from 2016	19.72
Treated from 2017	5.33

* This imply that they had their first selected SPP at the end of the previous year.

To include these schools in my analysis in this section I implement difference-in-differences approach with multiple periods and variation in treatment timing, by estimating the following equation.

$$Y_{it} = \beta_1 D_{it} + \alpha X_{it} + \pi_t + \pi_i + \varepsilon_{it} \tag{4}$$

Where Y_{it} corresponds to the outcomes I am going to evaluate; D_{it} is what is known as a policy variable, that is equal to one if a school, i, is treated at time t and zero otherwise; D_{i2015} is always zero for never treated schools. π_t are year fixed effects, π_i are school fixed effects and X_{it} is a vector of observed characteristics and ε_{it} are standard errors clustered at the level of the school. Once again, to refine the control group and make it more comparable, I use as a control those schools that could have had selected students in 2015 if the government had not changed the merit-based

[18] For instance, the correlation of the number of SPP in 2014 and 2015 is 75%.

criterion of 2014 and that could have had selected students in 2016 if the government had not changed the merit-based criterion of 2015. Nevertheless, I find similar results when relaxing the criterion for refining control group.

Tables A.X and A.XI. present the results for this model that show no effect for SPP on Saber 11 scores, but robust negative relation of having a selected SPP the previous year and the tests scores for Saber 9, Saber 5 and the dropout rate. Supporting the hypothesis of a change in the composition of skills. Although I do not interpret these results as a causal effect.

Table A.V. Difference in differences with different treatment timing

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D1 - Dummy if school had SPP	0.0130	-0.0159	-0.0630***	-0.0189*	-0.0231*	-0.0346***
the previous year	(0.0108)	(0.0105)	(0.0134)	(0.0113)	(0.0140)	(0.0125)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,062	23,062	23,062	23,062	23,062	23,062
R-squared	0.884	0.899	0.858	0.881	0.794	0.842

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.VI. Difference in differences effect of SPP on dropout rate (asset ownership as covariate)

	General	6th grade	7th grade	8th grade	9th grade	10th grade	11th grade
D1 - Dummy if school	-0.00192***	-0.00300**	-0.00200*	-0.00205*	-0.000733	-0.00256**	-5.47e-05
had SPP the previous year	(0.000736)	(0.00123)	(0.00110)	(0.00108)	(0.00107)	(0.00129)	(0.000858)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,032	23,032	23,032	23,032	23,032	23,032	23,032
R-squared	0.521	0.543	0.470	0.451	0.408	0.430	0.302

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix IV. Criterion for refined control group only applied for 2015

Table A.VII. Difference in differences effect of SPP on standardized test results (criterion for refined control group only applied for 2015)

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D^1_{i2015} - Dummy if school had SPP	0.0559***	-0.00536	-0.105***	-0.0516***	-0.0613***	-0.0509***
the first year of the program	(0.0154)	(0.0153)	(0.0185)	(0.0163)	(0.0201)	(0.0175)
D^2_{i2015} - Number of SPP	0.00613**	-0.00126	-0.0107***	-0.00304**	-0.00633***	-0.00367**
the first year of the program	(0.00241)	(0.00149)	(0.00199)	(0.00153)	(0.00196)	(0.00166)
D^3_{i2015}- Proportion of SPP	0.458**	0.183	-0.552**	-0.674**	-0.271	-0.0285
the first year of the program	(0.231)	(0.188)	(0.260)	(0.262)	(0.273)	(0.204)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,505	17,505	17,505	17,505	17,505	17,505

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.VIII. Difference in differences effect of SPP on dropout rate (criterion for refined control group only applied for 2015)

	General	6th grade	7th grade	8th grade	9th grade	10th grade	11th grade
D^1_{i2015} - Dummy if school had SPP	-0.00568***	-0.00820***	-0.00480***	-0.00684***	-0.00329*	-0.00454**	-0.000120
the first year of the program	(0.00111)	(0.00200)	(0.00176)	(0.00172)	(0.00173)	(0.00225)	(0.00151)
D^2_{i2015} - Number of SPP	-0.000388***	-0.000484***	-0.000383**	-0.000496***	-0.000353**	-0.000381**	-0.0000492
the first year of the program	(0.000105)	(0.000160)	(0.000157)	(0.000161)	(0.000147)	(0.000191)	(8.84e-05)
D^3_{i2015} - Proportion of SPP	-0.0444***	-0.101***	-0.0670***	-0.0583***	-0.0429***	-0.0580***	-0.00264
the first year of the program	(0.00942)	(0.0161)	(0.0162)	(0.0175)	(0.0151)	(0.0169)	(0.0115)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17,426	17,426	17,426	17,426	17,426	17,426	17,426
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix V. Using the constructed variable of asset ownership as covariate

Table A.IX. Difference in differences effect of SPP on standardized test results (asset ownership as covariate)

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D^1_{i2015}	0.0648***	0.0231	-0.116***	-0.0531***	-0.0518**	-0.0498**
	(0.0200)	(0.0199)	(0.0218)	(0.0202)	(0.0248)	(0.0213)
Average number	0.0824***	0.0555***	-0.0327***	0.00103	0.0158	0.000901
of assests per family	(0.0133)	(0.0117)	(0.0126)	(0.0107)	(0.0132)	(0.0120)
Average Age	-0.0233	-0.00720	-0.0532***	-0.0508***	0.0450**	0.0295*
	(0.0150)	(0.0141)	(0.0179)	(0.0156)	(0.0177)	(0.0158)
Women proportion	-0.629***	0.201	0.0651	1.193***	0.446**	0.625***
	(0.214)	(0.185)	(0.201)	(0.188)	(0.227)	(0.215)
Victims Proportion	-0.117	0.000445	-0.0345	-0.0704	0.180	0.204
	(0.138)	(0.101)	(0.132)	(0.116)	(0.155)	(0.146)
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,508	15,508	15,508	15,508	15,508	15,508
R-squared	0.891	0.903	0.865	0.887	0.799	0.847

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.X. Difference in differences effect of SPP on standardized test results (asset ownership as covariate) - Intensive margin treatments

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language
D^2_{i2015}	0.00605**	0.000222	-0.00993***	-0.00231	-0.00493**	-0.00298*
	(0.00248)	(0.00157)	(0.00199)	(0.00154)	(0.00203)	(0.00169)
D^3_{i2015}	0.492**	0.389**	-0.432	-0.611**	-0.0369	0.128
	(0.249)	(0.198)	(0.283)	(0.287)	(0.289)	(0.217)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,508	15,508	15,508	15,508	15,508	15,508

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

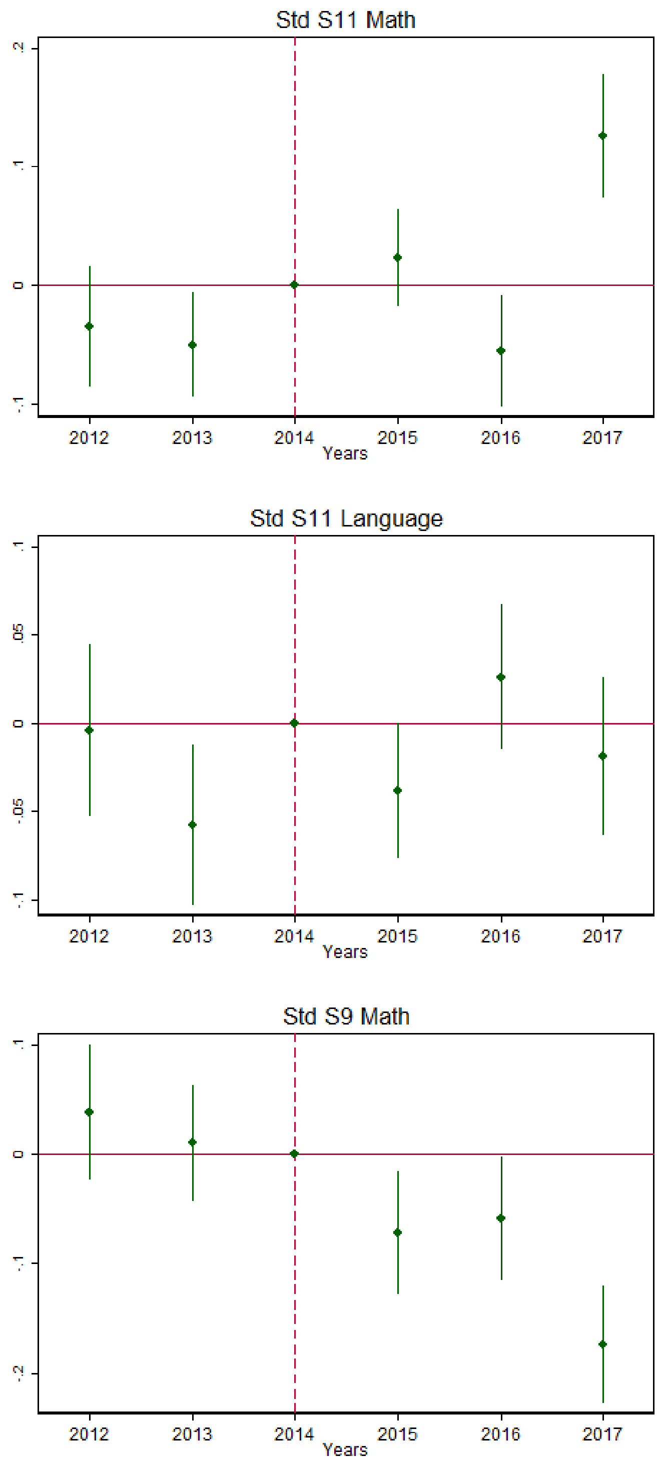
Table A.XI. Difference in differences effect of SPP on dropout rate (asset ownership as covariate)

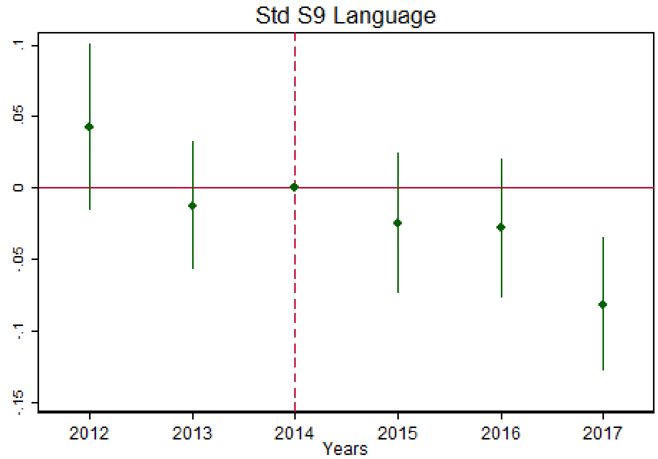
	General	6th grade	7th grade	8th grade	9th grade	10th grade	11th grade
D^1_{i2015}	-0.00564***	-0.00797***	-0.00448**	-0.00636***	-0.00316*	-0.00509**	-0.000811
	(0.00112)	(0.00200)	(0.00175)	(0.00172)	(0.00172)	(0.00248)	(0.00147)
D^2_{i2015}	-0.000378***	-0.000462***	-0.000364**	-0.000460***	-0.000340**	-0.000402**	-6.62e-05
	(0.000105)	(0.000160)	(0.000158)	(0.000160)	(0.000146)	(0.000194)	(8.72e-05)
D^3_{i2015}	-0.0445***	-0.100***	-0.0648***	-0.0529***	-0.0415***	-0.0585***	-0.00593
	(0.00930)	(0.0163)	(0.0156)	(0.0163)	(0.0150)	(0.0177)	(0.0117)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,438	15,438	15,438	15,438	15,438	15,438	15,438

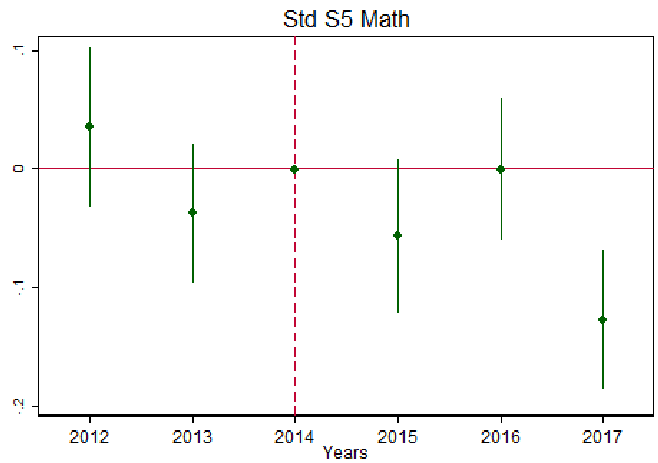
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

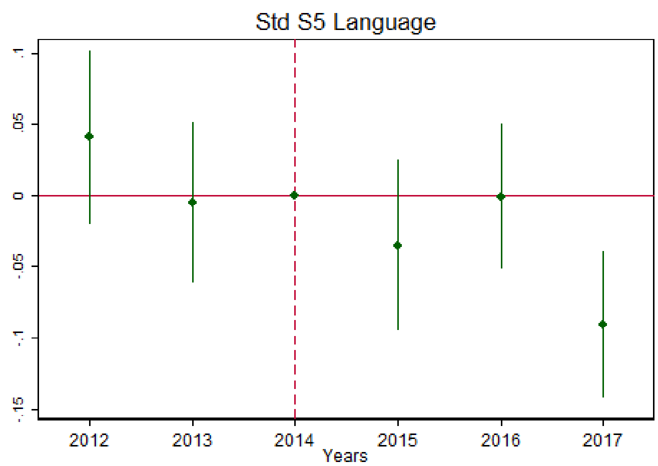
Appendix VI. Assessing parallel trends assumption

Figure A.III. Regression coefficients - interaction term (treated and year dummys)




Std S9 Language
.1
.05
0
-.05
-.1
-.15
2012
2013
2014
2015
2016
2017
Years


Std S5 Math
.1
0
-.1
-.2
2012
2013
2014
2015
2016
2017
Years


Std S5 Language
.1
.05
0
-.05
-.1
-.15
2012
2013
2014
2015
2016
2017
Years

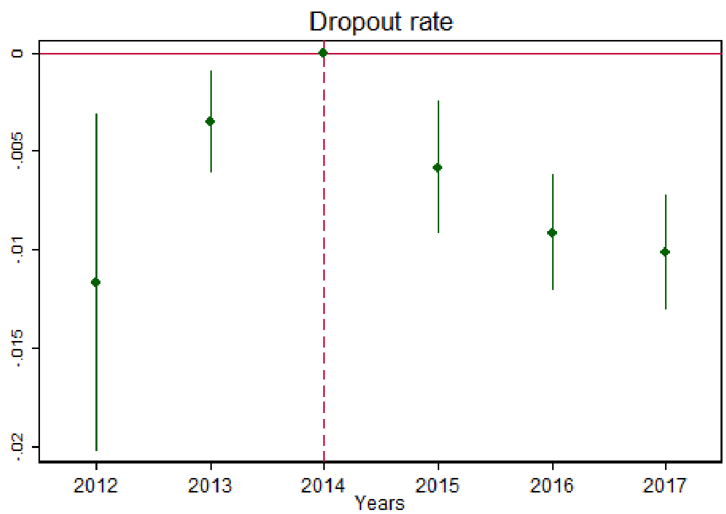


Table A.XII. Assessing parallel trends assumption for treatment 2 - Number of SPP in the first year of the program

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language	Dropout rate
D_{2012}	-0.00996***	0.000962	0.00206	0.00241	0.00574**	0.00330	-0.000640**
	(0.00387)	(0.00242)	(0.00237)	(0.00244)	(0.00280)	(0.00266)	(0.000319)
D_{2013}	-0.00906***	-0.00276	0.000602	0.00152	0.00278	0.00368	-0.000239
	(0.00278)	(0.00196)	(0.00233)	(0.00199)	(0.00251)	(0.00268)	(0.000254)
D_{2014}	-	-	-	-	-	-	-
D_{2015}	-0.000600	-0.00383**	-0.00462*	-0.000499	7.76e-05	0.000611	-0.000275
	(0.00180)	(0.00176)	(0.00243)	(0.00206)	(0.00275)	(0.00286)	(0.000246)
D_{2016}	-0.00733***	0.00135	-0.00518*	-0.00228	-0.00126	0.000386	-0.000655***
	(0.00250)	(0.00179)	(0.00270)	(0.00215)	(0.00314)	(0.00310)	(0.000240)
D_{2017}	0.00544*	-0.000837	-0.0164***	-0.000425	-0.00565**	-0.00282	-0.000893***
	(0.00278)	(0.00199)	(0.00289)	(0.00214)	(0.00269)	(0.00260)	(0.000244)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,567	15,567	15,567	15,567	15,567	15,567	15,567
R-squared	0.891	0.902	0.865	0.887	0.799	0.847	0.516

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.XIII. Assessing parallel trends assumption for treatment 3 - Proportion of SPP in the first year of the program

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language	Dropout rate
D_{2012}	-2.300***	-1.006***	0.385	-0.395	-0.688	0.508	-0.0407
	(0.452)	(0.371)	(0.398)	(0.398)	(0.445)	(0.470)	(0.0251)
D_{2013}	-2.452***	-1.039***	-0.385	-0.344	-0.286	0.213	-0.0109
	(0.328)	(0.306)	(0.338)	(0.325)	(0.354)	(0.324)	(0.0139)
D_{2014}	-	-	-	-	-	-	-
D_{2015}	-0.566**	-0.530**	-0.297	-1.062**	-0.146	0.541	-0.0608***
	(0.270)	(0.260)	(0.457)	(0.435)	(0.436)	(0.376)	(0.0163)
D_{2016}	-1.593***	0.235	-0.392	-0.573*	-0.375	0.199	-0.0587***
	(0.273)	(0.254)	(0.345)	(0.322)	(0.376)	(0.298)	(0.0146)
D_{2017}	-0.750**	-0.655**	-0.673*	-0.869**	-0.536	0.207	-0.0512***
	(0.333)	(0.305)	(0.374)	(0.346)	(0.380)	(0.328)	(0.0165)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,567	15,567	15,567	15,567	15,567	15,567	15,567
R-squared	0.892	0.902	0.864	0.887	0.799	0.847	0.516

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix VII. Entropy balance results

To provide further evidence on the conclusion regarding the effect of SPP on education achievement outcomes at the school level, I rebalance the sample to create groups that are more comparable in the pre-treatment period. To accomplish that I use a multivariate reweighting method presented in Hainmueller (2012) to reweight control group data to match the covariate moments in the treatment group. This entropy balance matching strategy attempts to reduce the imbalance, controlling for the confounding influence of pre-treatment variables.

I match schools on Wheter it is located in rural or urban area, whether the school is public or not and the constructed variable for students ownership of assets. Moreover I match on the outcome variable in the pre-treatment period to control for difference in pre-treatment trends as used for instance in (Collonelli and Prem, 2017). This technique would allow me to build better counterfactual data among never treated schools. After the matching process, I estimate the same model described in section 5 controlling by women ratio, victims of the conflict ratio, average age, and average socioeconomic strata, presenting these results in Table A.XIV.

Table A.XIV. Entropy balance results

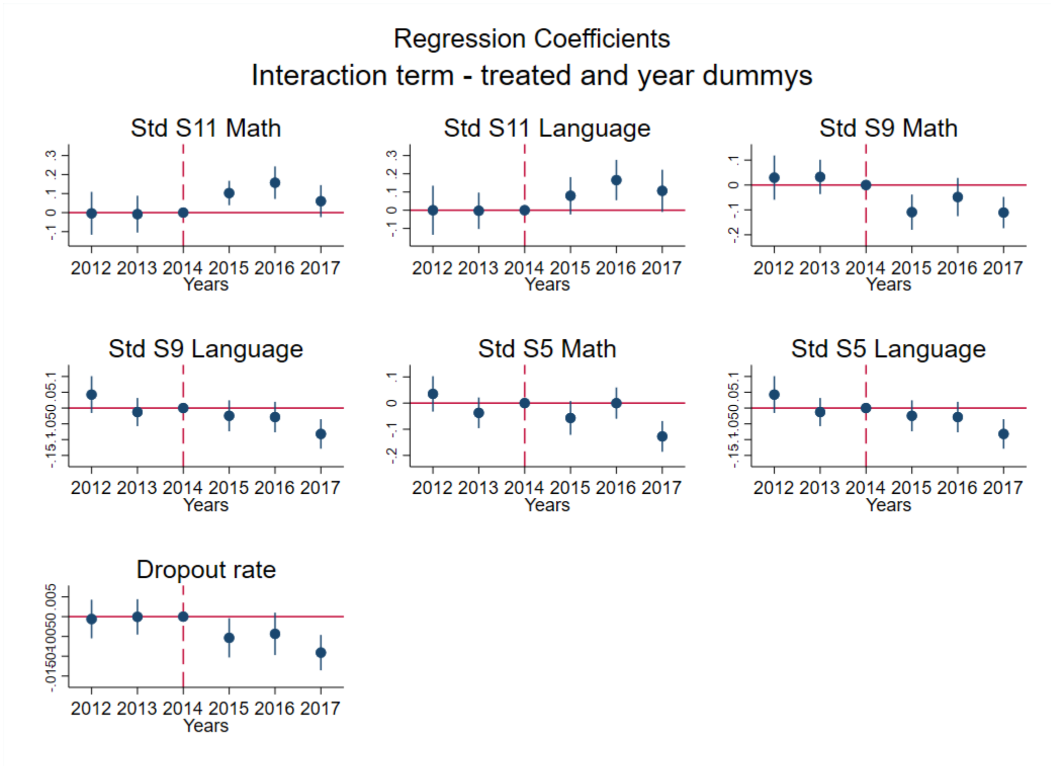
Treated units:	9768	total of weights:	9768
Control units:	959	total of weights:	9768

Before: without weighting						
	Treat			Control		
	mean	variance	skewness	mean	variance	skewness
Assets	2.84	0.53	0.26	3.15	1.27	0.13
Rural	0.10	0.07	2.89	0.25	0.17	1.20
Public	0.82	0.14	-1.71	0.53	0.25	-0.12
After						
	Treat			Control		
	mean	variance	skewness	mean	variance	skewness
Assets	2.84	0.53	0.26	2.84	0.83	0.29
Rural	0.10	0.07	2.89	0.10	0.08	2.77
Public	0.82	0.14	-1.71	0.82	0.14	-1.70

	(1) Std S11 Math	(2) Std S11 Language	(4) Std S9 Math	(5) Std S9 Language	(6) Std S5 Math	(7) Std S5 Language	(7) tasa_desercion
D^1_{i2015} - 1 if school had SPP	0.111***	0.118*	-0.107***	-0.0637**	-0.0542	-0.0844***	-0.00602***
the first year of the program	(0.0338)	(0.0262)	(0.0245)	(0.0270)	(0.0330)	(0.0270)	(0.00163)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

I also present the graphical results of the estimations when using the model described in equation (2) finding a balance in pre-treatment behavior now confirming the negative effect of having one SPP student for standardized tests results in both language and math in S5 and for math in the case of S9. I also find a positive effect of having a Selected SPP student on S11 and a negative effect on the dropout rate.



Appendix VIII. Heterogeneous effects - proportion of need-eligible students

Table A.XIV. Heterogeneous effects - proportion of need-eligible students

	Std S11 Math	Std S11 Language	Std S9 Math	Std S9 Language	Std S5 Math	Std S5 Language	Dropout rate
D_{i2015}	-0.161***	-0.162***	-0.129***	-0.0969***	-0.120***	-0.0760**	-0.00169
	-0.0293	-0.0246	-0.0364	-0.032	-0.04	-0.0368	-0.00155
Need-eligible	-0.0255	0.247***	0.240***	0.112*	0.213***	0.176**	-0.0131***
	-0.0579	-0.0572	-0.0776	-0.0649	-0.0826	-0.0748	-0.00477
D_{i2015}*Need-eli	0.339***	0.248***	0.0544	0.0926***	0.103**	0.0489	-0.00386**
	-0.0325	-0.0272	-0.0395	-0.0352	-0.0445	-0.0428	-0.00172
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,671	9,671	9,671	9,671	9,671	9,671	9,671
R-squared	0.917	0.922	0.86	0.888	0.803	0.847	0.596

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1