INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 02109

March 2022

Conditional Cash Transfers and High School Attainment

Evidence from a Large-Scale Program in the Dominican Republic

Manuel A. Hernandez

Jose A. Pellerano

Gonzalo E. Sanchez

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

AUTHORS

Manuel A. Hernandez (m.a.hernandez@cgiar.org), Senior Reseach Fellow at Markets, Trade, and Instittutions Division of the International Food Policy Research Institute (IFPRI), Washington, DC.

Jose A. Pellerano* (japellerano@siuben.gob.do), Senior Researcher at the Economic Analysis Department of Sistema Único de Beneficiarios (SIUBEN), Dominican Republic.

Gonzalo E. Sanchez (edsanche@espol.edu.ec), Associate Professor at Facultad de Ciencias Sociales y Humanísticas of the Escuela Superior Politécnica del Litoral (ESPOL), Ecuador.

* *Corresponding author*

Abstract

Conditional cash transfer (CCT) programs are widely implemented in developing countries but evidence of their medium- and long-term effects on educational achievements is still relatively scarce. This paper examines the impact of a large-scale CCT program on high school attainment in the Dominican Republic. We implement a quasi-experimental approach combining extensive educational, administrative, and household records from program participants across the country and exploiting variations in the scheme (amount) of school transfers received among program participants. We find that receiving additional transfers specific for high school education is, on average, associated with an 11.7-13.2 percentage points higher probability of completing high school relative to not receiving these transfers. We do not find major differences across urban and rural areas nor between female and male students. The transfers seem to play an important role during the last high school year of targeted students. The estimated impacts point to non-negligible effects on employment, salaries, and delayed parenthood. Several robustness checks support our findings.

Key words: High school attainment, conditional cash transfers, education, large-scale program, Dominican Republic

JEL codes: I28, I38, O15

Acknowledgements

We thank Jefrey Lizardo, Director of Sistema Unico de Beneficiarios (SIUBEN), for his support, especially during the final phase of the study. We also thank Dr. Ancell Scheker from the Ministry of Education for allowing accessing students' records. Similarly, we are very grateful for the support received at different stages of this work from Carlos Ariel Tejada, Carlos Julio Tejada, Olivia Bordeu, Rusbel Alcantara, Starling Sanchez, and Hernan Barahona, as well as collaborators from the Information Technology units at SIUBEN, Progresando con Solidaridad (PROSOLI), and Social Subsidies Administrator (ADESS).

1. Introduction

Over the past two decades, conditional cash transfer (CCT) programs have been extensively implemented worldwide –reaching over sixty countries– and have become one of the most popular social programs in developing regions. In Latin America, roughly one of every four individuals in 17 countries have received cash transfers, making up 20-25% of their household income (Robles, Rubio, and Stampini 2015; Ibarraran et al, 2017). These programs are mainly designed to both alleviate current poverty by providing direct income support to poor households and lessen future poverty by conditioning the transfers to compliance with certain education, health, and nutrition conditions aiming to improve the human capital of poor children and youth. While the literature on short-term effects of CCT programs on educational and non-educational outcomes is quite extensive (see, e.g., Fiszbein and Schady, 2009; Baird et al., 2014; Garcia and Saavedra, 2017; Ibarraran et al, 2017; Parker and Todd, 2017), much less is known about medium- and long-term effects of these programs in part due to the difficulty of following individuals over prolonged periods of time.

One relevant public policy question is whether CCT programs ultimately lead to higher human capital accumulation. In particular, whether the positive effects of these programs on school enrollment and attendance, which are generally observed in the short run, effectively lead to higher levels of educational attainment and school completion. We examine in this paper the impact of a large-scale CCT program in the Dominican Republic, Progresando Con Solidaridad (PROSOLI), on the completion of high school. We follow a quasi-experimental approach to assess the effect of receiving additional transfers specific for high school education on school attainment, relative to a comparable beneficiary group that do not receive these additional transfers.

Following the economic crisis in 2003, poverty and extreme poverty rates in the Dominican Republic spiked in 2004 reaching 49.5% and 15.4% (compared to 32.4% and 8.5% in 2002).[1] One of the policies the government took to address the situation was to implement a CCT program that started in October 2004 with the provision of conditional cash transfers for food to eligible socioeconomically vulnerable families, followed by cash transfers conditional on school enrollment and attendance in December 2005.[2] Since its establishment, the three main goals of

[1] Ministerio de Economía, Planificación y Desarrollo de la República Dominicana (2019).

[2] The cash transfers were accompanied with socio-educational programs and community services.

the program have been to improve the education, health, and nutrition of the targeted population, and the program has become one of the main social protection programs in the Dominican Republic with more than 800,000 conditional recipient households by the end of our analysis period in 2017, equivalent to almost 28% of the country's total population.[3] Regarding the average amounts transferred to families each month, these have been in the range of 12-19% of the country's minimum wage (depending on firm size) and 2.4-11.1% of the national food basket (depending on the household consumption quintile).

Until 2013, the only cash transfer for school enrollment and attendance was Incentivo a la Asistencia Escolar (ILAE). Households in the poorest target groups with children of school age were eligible for ILAE, ranging from elementary to high school. In June of 2013, the Bono Estudiantil Estudiando Progreso (BEEP) was created specifically targeting children in high school. For a household with one student in high school, BEEP transferred between 1.7 and up to 3.3 times more cash relative to ILAE depending on the grade level and high school format (regular or technical). The emphasis on high school attendance is very important. Approximately one in two male students and one in three female students in the public school system of the Dominican Republic who finish elementary school do not complete high school, despite the important returns of secondary education.

After several years of the program implementation, it is pertinent to evaluate the impact of the program on high school completion. We use extensive data from program participants and implement regression-based matching and subclassification (blocking) methods to assess the impact of BEEP transfers on high school attainment. We find that receiving additional transfers specific for high school education is, on average, associated with an 11.7-13.2 percentage points higher probability of completing high school compared to not receiving these transfers. We do not observe major differences on the effect of BEEP transfers across urban and rural areas nor between female and male students. The results are robust to alternative estimation methods and the use of different samples. Additional estimations point to the importance of providing BEEP transfers during the last high school year among targeted students. The impacts of these

---

Additional cash transfers (subsidies) and program components were introduced in the following years, which are described in the next section.

[3] The total number of recipients temporally increased to 1.56 million households during 2020 due to the COVID-19 pandemic and leveled off at 1.26 million by March 2021.

additional transfers on school performance in terms of grades and graduation on time are, however, mixed across urban and rural areas. Back of the envelope calculations indicate non-negligible effects of receiving additional school transfers on employment, salaries, and delayed parenthood.

The study contributes to the still relatively scarce literature analyzing medium- and long-term schooling effects of cash transfer programs. Two early studies are Behrman, Parker, and Todd (2005) and Behrman, Parker, and Todd (2011) that use differences-in-differences (DID) and matching methods to evaluate the impact of the Mexican CCT program Progresa/Oportunidades on schooling, more than five years after its implementation.[4] The former finds that there is a statistically significant difference of about 0.2 additional years of schooling between the group that was treated for 5.5 years versus the control group (treated for four years); the latter considers the same time frame as the previous work but accounts for both short (1-2 years) and long (5-6 years) treatment periods and shows that a longer program exposure has a larger impact on schooling (one additional year of schooling versus 0.2 years). More recently, Kugler and Rojas (2018) analyze the effects of Progresa/Oportunidades up to 17 years after its start and find that a youth exposed to seven years of the program has, on average, three additional years of schooling, 18 percentage points higher probability of completing high school, and five percentage points higher probability of graduating from college, relative to someone never exposed to the program. Parker and Vogl (2018), in turn, compare those individuals who were treated before turning 12 years old with those who were treated after the age of 15 (i.e., past the critical transition between primary and secondary school) and find that the group with the longest exposure completed, on average, 1.4 more years of school.

Outside of Mexico, Barham, Macours and Maluccio (2017) estimate the long-term educational effects (10 years after implementation) of a randomized phase-in CCT program in Nicaragua and demonstrate that early exposure to the program lead to around 0.3 additional years in school compared to late exposure. Araujo, Bosch, and Schady (2019) use a regression discontinuity design (RDD) approach to assess the effects of an unconditional cash transfer program in Ecuador and show that 10 years after the beginning of the program the effect on the

[4] As noted by Parker and Todd (2017), Progresa/Oportunidades is probably the most studied social program in a developing country, in part due to its emphasis on human-capital accumulation and because the evaluation data were made available relatively fast.

probability of finishing secondary school is between one to two percentage points higher than the counterfactual school completion rate.[5] In Colombia, Barrera-Osorio, Linden, and Saavedra (2019) explore the medium- and long-term effects (8-12 years after initial receipt) of three models of CCT programs assigned randomly to high-school students with scarce resources: transfers conditional on enrollment, transfers with obligatory savings, and transfers with additional incentives for high school graduation and college enrollment. While the effects of the three alternative payment structures on secondary graduation (i.e., taking the secondary school exit examination) are mixed relative to a control group that do not receive transfers, the program with obligatory savings has the largest effects in the long term and increases (high-quality) college enrollment rates by about 2.8 percentage points.[6] In Indonesia, Cahyadi et al. (2020) examine the cumulative impact (six years after) of a CCT program targeting poor households with children and lactating women across randomly selected subdistricts and find a four percentage points increase in school enrollment among children aged 7-15.

The paper also adds to the discussion in the Dominican Republic about the lasting effects of this large-scale social program, which has benefited more than a quarter of the population. Previous studies in the country have mainly focused on short-term educational effects and find mixed results. Lozano (2012), for example, finds that recipients of education transfers (ILAE) have a 2% lower probability of repeating the school year relative to nonrecipients, but he also finds a 4.5% higher likelihood of dropping high school. Reyes (2014) finds that ILAE participation increases the probability of overall school attendance by around 4%, while the probability of dropping school is reduced by about 1%. These studies are limited to cross-sectional data collected in 2010.[7] Our paper is the first study to assess the completion of

---

[5] Although the transfers were not conditional, households were still encouraged to spend the transfers on their children (see also Schady and Araujo, 2008).

[6] The program incentivizing high school graduation and college enrollment show similar effects on tertiary enrollment than the compulsory savings program; yet it only increases enrollment in low-quality colleges as opposed to the savings program that increases enrollment in four-year universities.

[7] Most of the previous studies use the Encuesta de Evaluacion de la Proteccion Social (EEPS or Social Protection Evaluation Survey) exclusively designed and implemented by the government in 2010 to evaluate the program. The survey includes 1,451 beneficiary households and 1,345 control households that share similar socioeconomic conditions (quality of life index) and live in the same neighborhood

secondary education among teens in beneficiary households using various administrative data for multiple years and exploiting variations in the scheme (amount) of school transfers received.

In the same vein, an additional novelty of the paper is the combination of three sources of extensive administrative data. The first source of information are the students' educational records from the National Exams (Pruebas Nacionales or PPNN) of the Ministry of Education. The second source is the Social Subsidies Administrator (Administradora de Subsidios Sociales or ADESS), which contains historical information of cash transfers received by the student's household based on the identification card of the household head. The third source is the Unique System of Beneficiaries (Sistema Unico de Beneficiarios or SIUBEN), which contains socioeconomic information of beneficiary (and non-beneficiary) households, before and during their participation in the program.

The remainder of the paper is organized as follows. Section 2 provides further details about the institutional background and description of the evaluated CCT program. Section 3 describes the data and methodology used for the analysis. Section 4 presents and discusses the estimation results. Section 5 concludes and provides policy recommendations.

2. Program description

The economic crisis prompted by the bankruptcy of three important banks in 2003 caused a 1.3% fall of the Gross Domestic Product (GDP), interrupting more than a decade of sustained growth in the Dominican Republic.[8] By February 2004, the depreciation of the Dominican peso and inflation reached annual rates of 104% and 61.4%, while the poverty and extreme poverty rate increased by around 17 and seven percentage points between 2002 and 2004, respectively. With the aim of reducing the negative impacts on the most vulnerable, the government launched the CCT program SOLIDARIDAD in October 2004, which was then renamed Progresando Con Solidaridad (PROSOLI) after its merge with the PROGRESANDO program in August 2012.[9]

(barrio) as the beneficiaries. Besides educational outcomes, Breton (2012) analyze the effect of the program on food consumption while Canavire and Vasquez (2012) examine the effect on employment indicators.

[8] The annual GDP growth rate in the country averaged 5.7% over the period 1992-2002.

[9] The program was relaunched under its new name SUPERATE in January 2021.

As noted in PROSOLI's operational manual, the foremost goal of the program is to reduce the intergenerational transmission of poverty by supporting households' investment in education, health, and nutrition and promoting human capital accumulation through the development of skills and opportunities for work and life in general (PROSOLI, 2017). The implementation of the program entails the provision of multiple cash transfers combined with socio-educational accompaniment through house visits and linkage of families to local community programs and services.[10] The program focuses on the poorest neighborhoods based on the country poverty map and targets the poorest households within these areas. SIUBEN is the public institution responsible for identifying, registering, and classifying families across these targeted areas according to a quality-of-life index (ICV, for its Spanish acronym), a proxy means test exclusively designed for the program implementation.[11] The ICV classifies families in the SIUBEN database into four categories based on their poverty level (from ICV-I to ICV-IV). Prioritized households, eligible for conditional cash transfers, are those in the first two categories, ICV-I and ICV-II, with special attention to families with pregnant women, children 0-5 years old, and school-age children.

The conditional cash transfers include Comer es Primero (CEP), Incentivo a la Asistencia Escolar (ILAE), and Bono Estudiantil Estudiando Progreso (BEEP). As shown in Figure 1, CEP started with the program launch in October 2004 and is a cash transfer for food conditional on households complying with scheduled health checks, putting emphasis on pre- and post-natal exams, immunization programs, and routine checkups for children until the age of five. The beneficiary households of CEP receive about US$ 16.3 each month to supplement the purchase of food at stores affiliated to the Social Supply Network (RAS, for its Spanish acronym).

Similarly, recipient households with members in the five to twenty-one years old range are eligible to receive ILAE and BEEP. To comply with program conditions children must be

[10] Additional program components were added after the merge of SOLIDARIDAD with PROGRESANDO such as digital registration of family members, human formation and public awareness, environmental protection, and access to new technologies of information through Community Technological Centers.

[11] The ICV includes four main group of variables to approximate a household's poverty level: dwelling characteristics, access to public services, ownership of home appliances, and household sociodemographic characteristics.

enrolled in school and should not miss more than 20% of school days. Between 2005 and 2013, the only transfer for school assistance was ILAE, designed for students both at primary and secondary levels without distinctions.[12] In June 2013, BEEP was created specifically targeting high school students, particularly among families that were already receiving ILAE transfers. Households whose children graduate from primary school and enroll into high school became eligible to receive BEEP transfers. The annual payment schedule of both ILAE and BEEP is distributed in five cycles of two months each, excluding transfers for July and August, which correspond to summer vacations. Table 1 shows the payment amounts under each transfer scheme (ILAE in Panel A and BEEP in Panel B) relevant for our period of analysis.

As observed, BEEP implied an important increase in the amount transferred to selected families with teens in high school. For instance, holding constant all other transfers, while a household with one member in high school received from ILAE US$ 13.4 every two months; with BEEP, a family with one member at the secondary level received US$ 22.3, 31.3 or 44.7 every two months depending on the grade level and high school format (regular or technical). Overall, BEEP caused an important surge in cash transfers for beneficiary families with members enrolled in high school; on average, transfers in the example above represented 1.7 up to 3.3 times more cash compared to ILAE.[13]

In addition to conditional cash transfers,[14] recipient households are eligible for unconditional transfers. This is the case of Food Supplement for the Elderly (Suplemento Alimenticio)

---

[12] During our period of study, the education system in the Dominican Republic comprised eight years of primary education and four years of secondary education. Starting in 2015, the education system gradually transitioned to a new grade division of primary and secondary education; now, each level has six grades, instead of eight grades in the first level and four grades in the second level. The 2018-2019 school year was the last one with fourth grade as the senior year of high school.

[13] BEEP transfers were suddenly suspended in mid-2016 and reestablished in early 2019. We discuss below the implications of this suspension as it affected the last high school year of our study period.

[14] In 2008 a cash transfer for college attendance named Higher Education Incentive (IES, for its Spanish acronym) was also introduced. The goal of this transfer (US$ 10 each month) is to promote college assistance among PROSOLI beneficiary households. Yet, only a tiny share of program beneficiaries has been recipient of this transfer (0.47% in our main database for analysis).

introduced in August 2007, Bonogas introduced in September 2008, and Bonoluz introduced in December 2009. Families with members aged sixty-five years or older are eligible for the Food Supplement for the Elderly with a monthly transfer of US$ 9. Bonogas and Bonoluz work as subsidy schemes to support, respectively, households' expenditure in gas for cooking (US$ 5.2 per month) and payment of energy bills (from US$ 1 up to US$ 10 per month depending on households' energy consumption). Households classified as ICV-I through ICV-III are eligible for all unconditional transfers while households classified as ICV-IV are only eligible for Bonogas.

Despite households meeting the eligibility criteria for the conditional and unconditional transfers, the ultimate number of beneficiaries for each transfer and accompanying services depends on budget availability. Hence, it is regularly the case of households living in targeted areas and satisfying eligibility conditions that are not recipients of some program components or have not been included yet. This is mainly the case of school transfers and certain subsidies such as Bonoluz and Suplemento Alimenticio, as opposed to CEP and Bonogas that cover a larger fraction of eligible households.

In the case of additional school transfers, while BEEP mainly targets households receiving ILAE transfers, several of these eligible households are not BEEP recipients. Two additional limiting factors are incompleteness (underreporting) of the student database maintained by the Ministry of Education and insufficient resources for field verification of schooling status among all eligible students (according to SIUBEN database). These limitations in program implementation do not result though in systematic differences between eligible households that end up receiving additional transfers specific for high school education and eligible households that do not receive these extra transfers.[15] We exploit these variations in the reception of cash transfers among eligible households in targeted areas to evaluate the impact of receiving additional BEEP transfers on high school attainment, while controlling for the reception of other program transfers.

[15] Informal discussions with PROSOLI officials further confirmed that these shortcomings occur nationwide and are not constrained to specific locations or populations neither depend on particular factors.

3. Empirical approach

3.1 Data

We gather and combine information from three administrative datasets for the analysis. The first source is the Ministry of Education's National Exams (PPNN) database. Until 2015-2016 and 2018-2019 school years, students at primary and secondary levels, respectively, were required to take tests administered by the Ministry at the national level –in eighth grade of elementary school and in fourth grade of high school, according to the previous division of these two levels–.[16] This first dataset that ranges from 2005 through 2017 contains student's ID, school location, format and type (public or private), school grades, and test results of subjects assessed at National Exams and if the student was promoted or not.[17] From these data, we observe a 57% high school graduation rate among students that finished primary school between 2005 and 2013. By gender, there is a 61% graduation rate among females and 52.2% among males.

The second database is provided by the Social Subsidies Administrator (ADESS), the institution responsible for depositing transfers each month according to the recipients' payroll received from PROSOLI's Administration. We have access to all transfers made to each recipient (household head) since the beginning of the program (October 2004). These data allow us to confirm the exact date at which families started receiving each of the transfers and the amount of money wired in each transfer. Between October 2004 and December 2017, over US$ 2.3 billion were transferred to nearly one million households.

The third database is the census data on beneficiary and non-beneficiary households provided by SIUBEN. The mandate of the institution is to list households located in targeted poor areas, estimate their ICV level, and determine families' eligibility to the different transfers. This dataset contains information on households' location and sociodemographic characteristics, dwelling characteristics, access to public services, among others. We use two waves of SIUBEN census

[16] Starting in 2019-2020 school year, only tests for the last grade of high school are administered (i.e., at sixth grade of high school under the new grade level division).

[17] The subjects evaluated are Spanish, Math, Social Sciences, and Natural Sciences. For a student to approve, she or he requires a final grade of at least 70 points (on a scale 0-100), where the final grade is obtained as the sum of the school year's subject grade –with a weight of 0.7– and the corresponding result in the National Exams –with a weight of 0.3–. Students must pass all subjects to be promoted to the next education level.

data. The initial listing or first wave, collected between 2004 and 2010, has information on beneficiary and non-beneficiary households before the former started receiving transfers from PROSOLI, thus serving as the household baseline data.[18]

Given the extent of missing last names for households' children in the first wave, we use SIUBEN's second wave data, collected in 2011 and 2012, for the matching process between SIUBEN household dataset and the corresponding student's school records in the National Exams (PPNN) database. While ADESS transfer records can be directly matched to the SIUBEN household data through the household's head national ID, we have to develop an algorithm to match students from the PPNN database to their respective households due to the absence of a common identifier with SIUBEN. After matching students with a corresponding child in the household, we retrieved information of the household collected at baseline using the household's head national ID. We keep only students in households where the head was the same in both SIUBEN waves.[19]

The algorithm searches for students among family members in SIUBEN database based on their first and last names, sex, age, and province. Conditional that both province and sex in both datasets coincide and the age absolute difference is not greater than three years, the algorithm matches a student in the PPNN database to the closest household child in SIUBEN database provided that they meet a minimum (lower bound) matching score, which measures the similarity between the first and last names of the two observations. Individuals not matched during this first stage are considered in a second phase where province is not binding but all other conditions are satisfied, and the lower bound matching score is raised –see Appendix A for additional details–.[20] We also assess below the sensitivity of our results to alternative pairing restrictions.

---

[18] 73% of the households in the raw dataset were initially listed between 2004 and 2005, 19% between 2006 and 2008, and the remaining 8% between 2009 and 2010.

[19] This assumes that students matched in the second wave were also household members during the initial household listing. The period difference between the two listings (waves) is five to six years for most households.

[20] The algorithm was developed in Python using the *FuzzyWuzzy* library. We conducted additional visual validations on random matched samples after the first stage with the aim of setting stricter lower bound matching scores and reduce the incidence of false positives in the overall matching process.

Following the above (two-stage) criteria, the matching process identified 219,659 students in both SIUBEN's second data wave and PPNN database who completed primary school during the 2010-2013 period, which constitute our relevant sample group as BEEP transfers for high school education started in the 2013-2014 school year.[21] After searching for their corresponding household heads in SIUBEN's first data wave, we were able to retrieve household information at baseline for 137,073 individuals (62.4% of students in the first group), which comprises our initial working sample.

Table 2 shows the distribution of students in our working sample according to the bundle of transfers received by their households. As observed, 29.9% of the observations correspond to students in non-recipient households; that is, households located in areas targeted by PROSOLI but non-eligible based on their quality-of-life index (i.e., mainly families in ICV-IV group). Another 11.1% correspond to students in households that only received unconditional transfers (i.e., mainly families in ICV-III group). Among the students in households that additionally received conditional transfers, i.e., families in ICV-I and ICV-II groups, those in households that received CEP represent 12.6% of the sample (CEP+ group), while those that also received ILAE transfers (ILAE+ group) represent 13.1%, and those that received both ILAE and BEEP transfers (ILAE-BEEP+ group) represent 33.4%. The + sign following the conditional transfer acronym is used hereafter to emphasize that households could have received other transfers.[22]

Appendix Table B.1 reports, in turn, the average (and standard deviation in parentheses) of the quality-of-life index (ICV) measured at SIUBEN's baseline wave, by bundle of transfers received in urban (Panel A) and rural (Panel B) areas, for households with no matched students

[21] Considering that secondary education comprised four years during the period of study, targeted households with students that graduated from elementary school in 2010 could have been exposed to one year (or more) of BEEP transfers if the students were still in high school in the 2013-2014 school year (or beyond), while households with students that graduated from elementary school in 2011 could have been exposed to two years (or more) of BEEP transfers, and so forth. The results are not much sensitive to additionally including households with students that graduated from elementary school in 2009 as a small group of them were still in high school after 2013.

[22] In the case of CEP+ group, they could additionally have received unconditional transfers (e.g., Bonogas, Bonoluz, and/or the Food Supplement for the Elderly); in the case of ILAE+ and ILAE-BEEP+ groups, they received CEP transfers but could also have received unconditional transfers.

with the National Exams (PPNN) database versus households with matched students included in our working sample. We generally do not observe major differences in the ICV between households with unmatched and matched students within each group, especially among conditional beneficiaries, which suggests that our overall matching process does not introduce a selection bias that could affect our estimations.[23]

The ILAE-BEEP+ group corresponds to students in households that started to receive BEEP transfers when the student attended high school. Recall that prior to high school, the student's household enrolled in PROSOLI is only entitled to receive ILAE transfers, while during high school the household may receive BEEP transfers or may continue to receive ILAE transfers due to constraints in the program implementation discussed above. Overall, ILAE-BEEP+ represents the group of beneficiaries with the highest intensity of treatment in terms of the amount of cash transfers received by the household and constitutes our main treatment group of interest for the impact evaluation analysis below.

Table 3 presents summary statistics of the sample groups by area of residence for a subset of variables measured at SIUBEN's baseline wave as well as number of years as program beneficiaries, total transfers received by the household, and if the student completed high school. Panel A corresponds to students located in urban areas and Panel B in rural areas. As expected, the ICV is decreasing as we move across treatment groups and households receiving school transfers (ILAE+ and ILAE-BEEP+ recipients) are poorer than those only receiving CEP+. Accordingly, non-beneficiary and non-conditioned beneficiary households are relatively more educated, have fewer members, and a larger of their dwellings are connected to water and sewage systems (i.e., they have a toilet or latrine at home), compared to conditioned beneficiary

[23] By construction, there is a larger proportion of pairings among the ILAE-BEEP+ group as SIUBEN's baseline wave includes households listed since 2004 and we focus on households with students that graduated from elementary school since 2010. Appendix Figure B.1 further shows similar ICV distributions within our two most relevant beneficiary groups, ILAE+ and ILAE-BEEP+, between households with no matched students and households with matched students included in our working sample. Although not reported, it is important to remark that we also do not observe major differences in the ICV (measured at SIUBEN's second data wave) between households that we were able to retrieve information at SIUBEN's first data wave (i.e., households with the same head in both SIUBEN waves) and households that we were not able to retrieve this information.

households in the three other groups. With respect to individual characteristics, while sex and age are roughly balanced across groups, students in beneficiary households that received conditional cash transfers present slightly lower academic performance compared to students in the first two groups as measured by the standardized average score of primary education national tests.[24] Similarly, households that received at least one of the school transfers, have been recipients for a longer period of time (10-10.4 years on average) and received a larger amount of transfers. The average annual amount received across areas by a non-conditioned beneficiary household is US$ 119-127 versus 231-242 for CEP+ beneficiaries, 336-340 for ILAE+ beneficiaries, and 402-407 for ILAE-BEEP+ beneficiaries. Finally, students in the ILAE-BEEP+ group exhibit the largest high school graduation rate (66-67%), similar to non-recipients.

3.2 Methodology

The estimation method follows a quasi-experimental approach as participation in PROSOLI is not random, but geographically targeted and household eligibility is determined by a quality-of-life index (ICV), a proxy means test calculated by SIUBEN. We implement both a regression-adjusted matching estimator and a blocking (also known as subclassification or stratification) estimator, which belong to the doubly robust type of estimators. The main advantage of these estimators is that they combine the estimation of an outcome regression model with a treatment model where only one of the two models need to be correctly specified to obtain an unbiased effect estimator (Bang and Robins, 2005; Imbens and Rubin, 2015). The availability of the variables used to construct the ICV that explain treatment status, including participation in different program transfers, allows us to (at least) rely on a correct specification of the treatment variable.[25] We focus on estimating the average treatment effect of receiving additional cash

[24] The age corresponds to the age of the individual as of 2016-2017, which is the last school year of the study period. Average test scores correspond to the standardized difference from the mean across the four subjects evaluated: Spanish, Math, Social Sciences, and Natural Sciences (in the first attempt).

[25] We alternatively considered implementing a Regression Discontinuity Design (RDD) to estimate treatment effects around the ICV cut-off point. The identifying assumption in a RDD is that the potential outcomes are continuous across the cut-off point, which implies that the density function of the variable that determines treatment (i.e., the running variable) is continuous at the threshold level. We tested this assumption using the method proposed by McCrary (2008) but we find statistically significant

transfers for high school education on the graduation of students among beneficiary households (hereafter the average treatment effect on the treated).

Matching and blocking estimators are based on the same conditional independence (ignorability) assumption or unconfoundedness. This assumption implies that, conditional on a set of pretreatment variables, the potential outcomes are orthogonal to the treatment status; that is, $\{Y_{0i}, Y_{1i}\} \perp D_i | X_i$, where Y_{0i} is the outcome of individual i in the absence of treatment, Y_{1i} is the outcome under treatment, D_i is the treatment variable and X_i is a vector of pretreatment (observable) variables. Given this assumption, conditioning on X_i eliminates selection bias, i.e. $E[Y_{0i}|X_i, D_i = 1] = E[Y_{0i}|X_i, D_i = 0]$, and the average treatment effect on the treated or $ATT \equiv E[Y_{1i} - Y_{0i}|D_i = 1]$ can be estimated by iterating expectations over X_i,

$$\delta_{ATT} = E\{E[Y_{1i}|X_i, D_i = 1] - E[Y_{0i}|X_i, D_i = 0]|D_i = 1\} \quad (1)$$

Following Rosenbaum and Rubin's (1983) propensity-score theorem, the ATT can be redefined as,

$$\delta_{ATT} = E\{E[Y_{1i}|p(X_i), D_i = 1] - E[Y_{0i}|p(X_i), D_i = 0]|D_i = 1\} \quad (2)$$

where $p(X_i) \equiv \Pr[D_i = 1|X_i]$ is the conditional probability of treatment or propensity score.[26] The other assumption needed to identify the treatment effect is the common support assumption or overlapping condition $c < p(X_i) < 1 - c$, for some $c > 0$, which ensures that observations in the treatment group have alike observations in the non-treatment group against which to compare them.

We calculate the ATT implementing matching and blocking estimators using propensity scores, augmented by regression. The matching and blocking ensure comparability between the treatment and comparison (control) group, while the regression adjustment helps for additional

---

discontinuities in the density function at the cut-off point of the ICV for different specifications. Such estimates thus invalidate the use of a RDD for the intended analysis.

[26] The theorem indicates that if the conditional independence assumption $\{Y_{0i}, Y_{1i}\} \perp D_i | X_i$ holds, then $\{Y_{0i}, Y_{1i}\} \perp D_i | p(X_i)$. See also Angrist and Hahn (2004) and Angrist and Pischke (2009).

bias removal and increased accuracy through further adjustment in covariates (Imbens and Rubin, 2015).[27]

For the regression-adjusted matching estimator, each observation in the treatment group is paired with the closest observation in the control group in terms of the propensity score and the difference in the outcome variable is adjusted by the discrepancy in observables across the matched observations using a regression framework. The overall bias-corrected matching estimator for the ATT is obtained as the average of the adjusted differences that is given by,

$$\hat{\delta}_{bcm}^{ATT} = \frac{1}{N_1}\sum_{i:D_i=1}^{N_1}\{Y_i - \hat{Y}_{0i}\} \tag{3}$$

where N_1 is the number of matched treated observations and $\hat{Y}_{i0}$ is the predicted (regression-based) non-treatment outcome.[28]

In the case of the regression-adjusted blocking estimator, the sample is divided into strata or blocks corresponding to different ranges of values of the propensity score. While a common practice in the literature is to use an ad hoc number of blocks (e.g., 6-8), we follow instead a data-driven approach as in Imbens and Rubin (2015) to determine the number of blocks. The following ATT expression is calculated within each block j,

$$\hat{\delta}^{ATT}(j) = \frac{\sum_{i:B_i(j)=1} Y_{1i} \cdot D_i}{\sum_{i:B_i(j)=1} D_i} - \frac{\sum_{i:B_i(j)=1} Y_{0i} \cdot (1-D_i)}{\sum_{i:B_i(j)=1}(1-D_i)} \tag{4}$$

where $B_i(j)$ is an indicator variable equal to one if individual i belongs to block j, and zero otherwise. The ATT is estimated within each block by a regression that, in addition to the treatment indicator, includes the propensity score and covariates X_i. The overall regression-adjusted blocking estimator for the ATT is computed as the weighted average of the estimated

[27] Despite the bias-reduction properties of matching and blocking estimators, the differences between paired treatment and control observations persist when the matching is not exact; in the case of blocking, the propensity score is not necessarily constant within blocks such that a regression allows for additional bias reduction within pairs or blocks.

[28] See Abadie et al. (2004) and Abadie and Imbens (2011) for additional details on this matching estimator and its variance estimator.

effects in each block, where the weight is the share of the treatment group that each block represents in the treated sample that is defined as,

$$\hat{\delta}^{ATT}_{strat} = \sum_{j=1}^{J} \frac{N_t(j)}{N_t} \cdot \hat{\delta}^{ATT}(j).^{29} \quad (5)$$

3.2.1 Common support and balance

Table 4 outlines the set of covariates X_i used to calculate the propensity score for the students in our working sample. These include the variables that are used to construct the ICV as well as other relevant demographic and socioeconomic characteristics. All the categorical variables linked to the derivation of the ICV are disaggregated into dichotomous variables according to the number of categories in which the original variable is divided.[30] The other variables include age, education, and literacy of household head, if spouse present, household size, vehicle ownership, whether the household runs any type of business in the dwelling, student birth year, and school district fixed effects.[31] The indicator variables for student birth year permit to control for common cohort effects across areas as well as for time (year) differences in the exposure to the program.[32] The propensity score $p(X_i)$ is estimated using a logit specification and is obtained separately for urban and rural areas since the variables used to calculate the ICV and the treatment cut-off points differ between these areas.[33]

[29] The variance of the estimator is similarly derived as the weighted average of the estimated variance in each block, where the weights in this case are the squared shares of the treatment group that each block represents.

[30] For example, floor material is divided into five dichotomous variables: cement, granite, wood, dirt, and other floor material.

[31] The results are not sensitive to alternatively controlling for household province location (instead of school district).

[32] In the estimations below, we further account for whether the household received subsidies for the elderly (Suplemento Alimenticio) and electricity (Bonoluz) since we observe some variation, although small, across our treatment and control groups in the exposure to these unconditional transfers; the gas subsidy (Bonogas), in contrast, basically covered all households in the analysis.

[33] The full set of variables considered in urban areas are 75 and 66 in rural areas.

Appendix Figure B.2 reports the distribution of the estimated propensity scores when modeling the ILAE-BEEP+ group, which is our main treatment group of interest, against: non-beneficiaries (non-recipients located in targeted areas) and unconditional beneficiaries (in Panel A), CEP+ group (Panel B), and ILAE+ group (Panel C).[34] The graphed variable is the linearized propensity score equal to $\ln[p(X_i)/(1-p(X_i))]$. It is clear the larger overlapping of the distributions (range of common support) as we move across panels, which indicates that the ILAE-BEEP+ group is more comparable to the ILAE+ group, followed by the CEP+ group and, ultimately, to the non-beneficiary and unconditional beneficiary group. This pattern holds both for urban and rural areas and is certainly correlated with the program's targeting scheme of conditional and unconditional transfers based on households' socioeconomic status within the areas of intervention. Similar overlapping patterns are observed when comparing the ICV distributions across the corresponding groups.[35]

Our base results concentrate in comparing the high-school graduation rate of the ILAE-BEEP+ group relative to the ILAE+ group, which are the two most alike groups such that we can better control for potential selection bias. As noted earlier, BEEP entailed an important increase, compared to ILAE, in the amount of cash transfer received by households with members enrolled in secondary school, so the estimated effect should be interpreted as the net impact of receiving these additional cash transfers. For completeness, we similarly compare graduation rates between ILAE-BEEP+ and CEP+ groups and between ILAE+ and CEP+ groups, although the CEP+ group generally comprises fewer poor households and thereby is less comparable to the other two groups that receive school transfers in varying degrees. In all estimations, we impose a standard trimming of one percent and drop all observations below the first percentile of the propensity score of treated observations and all observations above the last percentile of control observations to ensure the common support assumption.[36]

[34] Non-recipients and beneficiaries not receiving conditional transfers are grouped together as their distribution of propensity scores is very similar.

[35] Details are available upon request.

[36] We opt for this procedure rather than limiting the working sample to, for example, above the 5 percentile and below the 95 percentile because we do not have few observations with propensity scores at the tails of the distributions for the corresponding treatment and control groups (that could otherwise affect our estimations). The results are though not sensitive to this alternative trimming.

Appendix Table B.2 provides additional insights about the balance between ILAE-BEEP+ and ILAE+ groups. The table reports the standardized mean differences between these two groups across all variables (measured at SIUBEN's baseline wave) that are used to calculate the propensity scores for all male students in urban and rural areas included in our main estimation sample.[37] Columns (1) and (3) correspond to the unmatched sample differences and columns (2) and (4) to the matched sample differences. We generally do not observe major differences across variables between the two groups; from columns (1) and (3), in only 7 cases the difference is greater than 0.1 in absolute value in urban areas and 10 cases in rural areas. Once ILAE-BEEP+ recipients are matched to their closest ILAE+ recipient based on their propensity score, the differences are further reduced (there are no cases where the absolute difference exceeds 0.04).[38]

When stratifying the sample into blocks using the subclassification estimator, the differences between the two groups are also very small, despite that the grouping of observations under this method is less granular than in the one-to-one matching. The absolute difference between the two groups is greater than 0.1 in 43 (6.2%) of the cases but never greater than 0.2 out of 693 comparisons across ten blocks in urban areas, and greater than 0.1 in 51 (10.2%) of the cases but greater than 0.2 in only 3 cases out of 498 comparisons across eight blocks in rural areas.

4. Results

This section presents and discusses the estimation results. Sections 4.1 through 4.3 focus on the comparison in high school graduation rates between students exposed to ILAE-BEEP+ school transfers, which is our main treatment group of interest, and students only exposed to ILAE+ transfers, which is the group that closest resembles the former. We separately estimate the ATT for urban and rural areas due to cutoff differences in program inclusion across areas of residence and differences in the variables used to construct the eligibility index model (quality-of-life

[37] The results are very similar for female students considering that most variables are calculated at the household level. After discretizing the variables, a small number of them are dropped in the propensity score estimation due to the lack of observations among the treatment and/or control group in urban or rural areas. In most cases, these correspond to residual category variables (e.g., other wall material).

[38] These minor differences are important for the subsequent regression adjustment performed as regression methods are not considered suitable for adjusting differences greater than 0.25 in absolute value (Imbens and Rubin, 2015).

index or ICV). The standard errors reported are heteroskedastic robust and the results are statistically significant at the 1% level unless stated otherwise. In Section 4.4, we present results comparing ILAE-BEEP+ and CEP+ recipients, as well as ILAE+ and CEP+ recipients. In Section 4.5, we analyze effects on school performance indicators, including grades and on-time graduation. Finally, in Section 4.6 we perform back of the envelope calculations to approximate eventual transfers' effects in the labor market and delayed parenthood per dollar invested.

4.1 Base results

Panel A in Table 5 presents the results including all female and male students. We report estimates using matching and blocking estimators and include both unadjusted and regression-adjusted ATT –our preferred estimate given their conditional bias reduction property. Note, however, that the magnitude of coefficients is somewhat similar both across and within estimators, which indicates that our results are generally not much affected by the choice of the estimator and suggests that the samples across the comparison groups are relatively well balanced.

Based on the adjusted estimates for the matching and blocking estimators, we do not observe major differences between urban and rural areas. We find that exposure to both school transfers increases the probability of high school completion by 11.7-13.2 percentage points (pp) in urban areas and by 12.4-12.9 pp in rural areas, compared to exposure to ILAE transfers only. Considering the average graduation rates among the matched control groups (reported in the last rows in each panel), the estimated effects represent an increase in the high school graduation rate of 23-25.3% compared to the counterfactual with no additional BEEP transfers.

The magnitude of these estimates is certainly important, considering that the effects are driven by households' exposure to additional transfers for high school attendance and net of other transfers, socio-educational support, and services received by both groups. While we do not have data to assess how these monetary transfers operate (or are allocated) within the household, BEEP transfers seem to be an important extra monetary incentive to motivate student household members to attend and graduate from high school and/or to encourage their parents to encourage them finishing school.

Panels B and C of Table 5 report estimates disaggregating by sex. We similarly do not find a clear difference on the impacts between female and male students, which vary by location. While the ATT is a bit higher for females compared to males in urban areas, the opposite is true in rural areas. The regression-adjusted estimated effect of receiving additional cash transfers is 11.4-12.4 pp for females versus 10.3-10.9 pp for males in urban areas but 9.7-10.9 versus 12.6-12.9 pp, respectively, in rural areas. These effects translate for females into a 19.8-21.4% increase in the graduation rate in urban areas and 16.1-18.1% increase in rural areas compared to the contrafactual with no BEEP transfers, while for males the estimated effects represent a 21.9-23.5% increase in graduation rates in urban areas and 27.8-28.5% increase in rural areas.

4.2 Timing of treatment

Our treatment group includes students that were exposed to BEEP transfers at different stages of their secondary education. As noted above, BEEP started during the 2013-2014 school year such that treatment students that graduated from elementary in 2010 were mainly recipients of BEEP transfers during their last high school year. Similarly, students graduating from elementary in 2011 or 2012 mostly received these additional transfers during their last two or three high school years, respectively. Finally, students graduating from elementary in 2013 were supposed to receive these additional transfers during all four high school years, but BEEP was suspended in June 2016 due to a sudden administrative decision by the Ministry of Education, which resulted in no BEEP transfers during their expected last high school year (2016-2017). We are thus interested in assessing whether there are differential effects on high school attainment by year of student graduation from elementary school.

Table 6 reports separate ATT estimates on high school graduation for students that completed primary level in 2010 through 2013 by area. Given the similar results obtained with the matching and subclassification methods, the analysis hereafter focuses on the matching (regression-adjusted) estimations. We find two interesting results. First, the effects for students that graduated from elementary school between 2010 and 2012 (Panels A-C) are very similar, ranging from 11.3 to 12.1 pp in urban areas and 13.5 to 15.2 pp in rural areas, which translate into higher secondary education completion rates of 20.8-22.2% and 24.9-29.2%, respectively. Yet, the effect for students that graduated from elementary school in 2013 (Panel D) is considerably lower: 7.1 pp in urban areas and 3.4 pp (marginally significant) in rural areas, which are, respectively, equivalent to a 15.7% and 7% increase in high school completion rate.

These findings suggest that the stage or timing at which the student is exposed to additional school transfers could play some role in high school attainment and could be even more relevant than the overall years receiving these transfers. Despite the average differences in the years of exposure to BEEP transfers across treatment students graduating from elementary in 2010 through 2012, they were all generally exposed to these transfers during their last high school year and we find similar effects across these three groups. In contrast, treatment students graduating from elementary school in 2013, were basically exposed to BEEP transfers during most of their high school except for the last year, when the transfers were abruptly suspended. This program holdup seems to have played a counteractive effect on high school graduation and points to the importance of receiving or at least to continue providing these additional transfers during the last school year among targeted students.[39]

It is also worth noting that households of most of the students in both our treatment and control group (over 96%) were enrolled into PROSOLI well before the student finished elementary school and were thereby recipients of different program benefits, including cash transfers for primary education, several years before starting secondary education. As reported earlier in Table 3, households of high school students receiving BEEP or ILAE transfers were, on average, enrolled into the program for over ten years. We find very similar results to our base results if we exclude students whose households were enrolled into the program after they graduated from primary school.[40]

4.3 Robustness checks

We now turn to assess the robustness of our results along two dimensions. The first dimension corresponds to sensitivity analyses around certain schooling characteristics and the algorithm used in the sample construction (reported in Appendix Table B.3), while the second dimension concerns robustness checks regarding the unconfoundedness assumption (reported in Appendix Table B.4). For ease of comparability, the last row in both tables shows the corresponding estimates for the base results (reported in Table 5).

---

[39] Certainly, we cannot fully discard time-varying unobservable differences between cohorts driving the observed results, but we are unaware of specific factors (other than the program suspension) that differentially affected treatment students graduating from elementary in 2013 during their last high school years, as opposed to treatment students graduating from elementary in 2010 through 2012.

[40] Details are available upon request.

4.3.1 Schooling features and sample construction

First, we consider potential confounding effects on high school graduation driven by the introduction of Extended School Day (ESD) program. In 2012, the Ministry of Education launched the ESD initiative with the inclusion of the first 21 public schools; by 2016-2017, the last school year in our period of analysis, 656 schools in our sample were operating under this scheme (from a total of 2,576 schools). As its name suggests, the main feature of this program is the extension of the school day from five to eight hours plus the expansion of the public-school nutrition program beyond breakfast by including a snack and lunch. These changes could positively influence high school attendance and completion, and thereby affect our results obtained considering the higher incidence of the ESD program among the ILAE-BEEP+ group relative to the ILAE+ group (24.1% versus 13.9%). We accordingly re-estimate our base model excluding students that attended schools participating in the ESD scheme. The estimation results are presented in Panel A of Appendix Table B.3. We observe that the ATT is 10.5 pp in urban areas and 11.6 pp in rural areas, which are only slightly lower than our base effects. This finding suggests that our base results are not mainly driven by potential confounding effects of the ESD program.

Second, we assess whether the school format might be confounded with the positive effects of receiving additional transfers on high school completion. In particular, the government generally put more emphasis on technical (vocational) high schools in recent years, including transforming several public schools' formats from regular to technical, which could have encouraged more students to complete high school under the latter format. Considering that BEEP started later than ILAE and a larger share of the ILAE-BEEP+ group (10.9%) attended technical schools relative to the ILAE+ group (6.9%), it is thus worth exploring the sensitivity of our results to excluding students attending technical schools. Panel B of the table shows the results when excluding students from technical high schools. We find that the ATT is 10.6 pp in urban areas and 12.3 pp in rural areas, which are roughly close to our base results, especially in rural areas. Hence, attending technical or vocational high schools does not seem to be confounded with the positive effects of receiving additional cash transfers.

In a related vein, in Panel C we drop students attending private schools, which comprise 10% of the ILAE-BEEP+ group and 7.3% of the ILAE+ group, as private schools are typically better equipped and dispose of more resources per student and the households of students attending these schools could further share specific socioeconomic characteristics that could overall influence the decision to attend and complete secondary education. Interestingly, we find somewhat larger effects than in the base estimates: 13.3 pp in urban areas and 14.2 pp in rural areas. These stronger effects when only considering students attending public schools could be correlated with the fact that the transfers become more relevant (salient) among students in public schools that, on average, have access to less resources both at school and home.

Fourth, as mentioned in Section 3.1 (and explained in detail in Appendix A), the matching algorithm between the SIUBEN and National Exams (PPNN) datasets comprises two phases: while in the first stage matching by province is binding, in the second stage it is not binding (reason why we raise the lower bound scores to match observations across datasets). Provided that the second-stage procedure could still be less accurate pairing observations than the first stage, Panel D of the table reports the results excluding students matched during the second stage –18.6% in the ILAE-BEEP+ group and 23.2% in the ILAE+ group–. The ATT is 11.7 pp in urban areas and 13.1 pp in rural areas, which are fairly close to our base results and provides additional support to our estimation approach.[41]

Lastly, in Panel E we report the results when restricting the absolute age difference to a maximum of two years (instead of three years) when pairing the SIUBEN and PPNN databases. The estimates are in this case 13.4 and 14 pp in urban and rural areas, respectively, which are slightly stronger but still close to our base effects.[42]

4.3.2 Unconfoundedness

We perform two standard robustness checks around the plausibility of the underlying conditional independence assumption for the estimators implemented above. The first test consists in

[41] Although not reported, we alternatively only included all paired observations that match in province in the first stage and are above the first decile (instead of the 25th percentile) of the matching score distribution, and we also find similar results to our base effects (11.3 pp in urban areas and 12.3 pp in rural areas).

[42] We identified certain cases where the student identification number changed between elementary and high school (2.1% of observations in ILAE-BEEP+ group and 2.0% in ILAE+ group). While these cases may be explained by a simple typo or coding error in the national databases, they could also imply a possible inaccuracy in our sample construction. The estimates, however, are not much affected when excluding these observations (12.9 pp in urban areas and 12.7 pp in rural areas).

estimating impacts on groups (or variables) that should not be affected by treatment or affected to a lesser extent than the main treatment group (or variables) –referred in the literature as pseudo treatment groups (or pseudo outcomes)–. Another sensitivity test consists in evaluating the effect of receiving treatment using a subset of the original pretreatment variables –i.e., subset unconfoundedness–. The intuition of the second test is that if the conditional independence assumption holds when controlling for a smaller number of confounders than all the control variables used in the analysis, then, it is more plausible that it will also hold when considering a larger (unobserved) number of variables.[43]

Panel A of Appendix Table B.4 presents the ATT using students in the CEP+ group as the treatment group and non-recipients and unconditional beneficiaries as the comparison group (that most closely resemble the CEP+ group). Since the former have not been exposed to school transfers, we would expect to find no statistically significant results in this case or at least smaller effects on high school graduation compared to our base estimates. As observed, the estimate for urban areas is not statistically different from zero at conventional levels, while the estimate for rural areas is small (2pp) and marginally significant at the 10% level.

Panel B shows, in turn, the estimates using a subset of the original pretreatment covariates to derive the propensity score and subsequent matching estimator. The covariates excluded are roof materials, lighting source, type of cooking fuel, and the additional demographic and socioeconomic variables that were not part of the original ICV model, resulting in a total of 27 excluded variables in urban areas and 24 excluded variables in rural areas –roughly one third of the initial pretreatment set of variables–. As observed, the estimates are 12.5 pp in urban areas and 12.7 pp in rural areas, which provide additional evidence supporting the identifying assumption.[44]

[43] As suggested by Imbens and Rubin (2015), it is still possible that the more restrictive assumption of subset unconfoundedness holds, while the assumption of unconfoundedness does not. In practice, however, this is less likely when qualitatively similar variables are available both in the original set and in the subset of pretreatment controls, as in the present exercise.

[44] We also tested excluding other subsets of variables (within the several dwelling characteristics, access to public services, home appliances, and household sociodemographic characteristics) and find similar results. Details are available upon request.

4.4 Other comparison groups

We are also interested in assessing the effects of school transfers on high school graduation relative to not receiving this kind of transfers among PROSOLI beneficiaries. The more immediate comparison group available for such exercise is the CEP+ group that only receive cash transfers for food conditional on households complying with scheduled health checks. Yet, while CEP+ recipients compose the group that more closely resembles the two groups receiving school transfers, we acknowledge that they are still not quite comparable due to the lower poverty prevalence at baseline among CEP+ households (reflected also in the lower range of common support in propensity scores between these groups). We thus separately examine graduation rates among the ILAE-BEEP+ and ILAE+ group against the CEP+ group with the associated caveats.

Appendix Table B.5 presents the regression-adjusted matching estimations comparing ILAE-BEEP+ versus CEP+ in Panel A and ILAE+ versus CEP+ in Panel B by area and sex. When comparing ILAE-BEEP+ against CEP+, we find a positive effect of being exposed to both school transfers of 5.6 pp in urban areas and 5.5 pp in rural areas −first row in Panel A− and the impact is slightly higher among female students −third row in Panel A−. In contrast, when comparing ILAE+ against CEP+, we find overall negative effects in both urban and rural areas (-4.7 and -4.9 pp, respectively) −first row in Panel B−. Although the results could be driven by the lesser comparability across groups (and different factors at play explaining high school graduation), they point to a key influence of receiving additional BEEP transfers (as opposed to ILAE transfers) on high school completion among beneficiary households.

4.5 Effects on school performance: Grades and on-time graduation

We now turn to examine the association between receiving additional school transfers and school performance indicators. We first assess the potential impact of receiving BEEP transfers on overall grades. We implement a standard maximum likelihood Heckman procedure where the selection equation models the probability of taking the national test at the end of secondary school and the main equation models the average (standardized) score obtained across the four evaluated subjects (Spanish, Math, Social Sciences, and Natural Sciences). We control for the same set of covariates as in our base estimations.[45]

[45] We accordingly exclude household size from the main equation to satisfy the exclusion restriction in

Appendix Table B.6 shows the estimation results by area and sex. We find mixed results across urban and rural areas. The reception of additional school transfers is positively correlated with grades in rural areas, but we find the opposite in urban areas (especially among male students); receiving BEEP transfers in rural areas is associated with 0.06-standard deviations score increase, compared to only receiving ILAE school transfers, while in urban areas it is associated with 0.05-standard deviations score decrease. Note that across all specifications, ILAE-BEEP+ recipients are also more likely to take national high school exams than ILAE+ recipients, which is in line with the positive effects on high school graduation discussed above.

Similarly, we assess whether receiving additional school transfers encourages people to graduate on time (i.e., spend four or less years to graduate from high school).[46] We estimate in this case a Probit model with sample selection where the selection equation models the probability of completing high school and the main equation models the probability of graduating on time. Appendix Table B.7 presents the estimation results by area and sex. While in urban areas we find a negative correlation between graduating on time and receiving BEEP transfers (compared to only receiving ILAE transfers), in rural areas we find a positive correlation; yet all correlations are not statistically significant except for female students in rural areas (significant at the 10% level).

Overall, receiving additional cash transfers from BEEP appears to have an important impact on high school graduation both in urban and rural areas and across female and male students. These transfers, however, do not necessarily also improve school performance in terms of grades and graduation on time, particularly in urban areas where we observe the opposite for grades (and no effects on graduation on time). This is certainly an aspect to be further explored in future work as more detailed data at the household and individual level is obtained.

---

this type of models, as the number of household members is generally associated with attending (dropping) school but not necessarily with better school performance, as observed in our case; see, e.g., Black et al. (2005), Maralani (2008), and Booth and Kee (2009) for a broader discussion on family size and children education. We also tested excluding other control variables (such as vehicle ownership and whether the household runs any type of business in the dwelling) from the main equation, but the results are qualitatively similar.

[46] The program did not establish though additional incentives or penalizations to finish high school earlier or later.

4.6 Cost-Benefit analysis

To put the results in perspective, we perform back of the envelope calculations to approximate the potential benefits of receiving additional BEEP transfers on employment, salaries, and delayed parenthood. This analysis is relevant in a context where the returns to education are not small and where the fertility rate among adolescents 15-19 years old is one of the highest in the Caribbean region (22%). Estimates from the 2017 National Workforce Survey (Encuesta Nacional Continua de Fuerza de Trabajo – ENCFT) indicate that the probability of employment for individuals in urban (rural) areas who complete high school is 79% (60%) higher than individuals who only complete primary school and 25% (19%) higher than individuals who start but do not complete high school; in addition, individuals who complete high school in urban (rural) areas earn US$ 72 (64) more a month than individuals with only primary education and US$ 30 (29) dollars more than individuals with incomplete secondary education.[47] Similarly, estimates from the 2013 National Demographic and Health Survey (Encuesta Demografica y de Salud – ENDESA) reveal that when controlling for socioeconomic (poverty) condition, adolescents in urban (rural) areas who complete high school have children, on average, 2.7 (2.4) years later than adolescents with only primary education and 1.8 (1.4) years later than adolescents with incomplete secondary education.

In Table 7 we combine these estimates with our base ATT estimates on completing high school using both propensity score matching and subclassification methods (reported in Table 5) and we normalize them by the average amount of additional transfers received per year by the ILAE-BEEP+ group, compared to the ILAE+ group (obtained from Table 3). We overall find non negligible effects of receiving additional BEEP transfers, relative to ILAE transfers, on employment, salaries, and delayed parenthood. For example, relative to someone who only completed primary education, for every additional US$ 10 transferred to the ILAE-BEEP+ group the student is expected to increase her/his chances of finding a job by 1.3-1.5% in urban areas and 1.2-1.3% in rural areas, while the returns on monthly salaries are US$ 1.2-1.3 in urban areas and US$ 1.3 in rural areas. Relative to someone who did not complete high school, the corresponding chances of finding a job are 0.4-0.5% and 0.4% higher and the returns on monthly

[47] The estimates on employment and salaries are obtained using Logistic and linear regression models, respectively, controlling for experience, experience squared, and month fixed effects. The full set of estimations are available upon request.

salaries are US$ 0.5-0.6 and US$ 0.6 higher. The delay in having a first child from every additional US$ 10 transferred is 0.6 months in both urban and rural areas, relative to completed primary education, and 0.4 and 0.3-0.4 months, relative to incomplete secondary education.

5. Concluding remarks

This paper examines the impact of a widespread CCT program on high school completion in the Dominican Republic. We focus on the effects of providing additional BEEP transfers specific for attending high school relative to only providing standard ILAE school transfers. We find that receiving these additional transfers are positively associated with high school completion both in urban and rural areas and among female and male students. The estimated average effects range between 11.7 and 13.2 percentage points, which are equivalent to an increase in the high school graduation rate of 23-25.3%. Complementary estimations suggest the importance of providing these additional transfers during the last high school year among targeted students. Regarding potential impacts on school performance indicators, we find mixed results in terms of grades and on-time graduation across urban and rural areas. Back of the envelope calculations reveal non-trivial effects of receiving additional school transfers on employment, salaries, and delayed parenthood.

This work contributes to the growing but still scant literature on medium- and long-term impacts of CCT programs in developing countries and is the first in the Dominican Republic to comprehensively evaluate the effect of this large-scale program on school attainment, more than a decade after its inception. The merge of three extensive administrative datasets allows us to implement several analyses around model identification and the construction of a plausible counterfactual, as well as to perform multiple robustness checks.

Our results also intend to contribute to the continuous discussion of an eventual redesign of the program, particularly after the substantial increase in the budget allocation for public education since 2012 –from 2% to 4% of the GDP– and subsequent discussions about the best use of these resources besides construction of new schools, rebuilding of old schools, and the need of incentive schemes to improve teachers' aptitudes and abilities. Given the high rates of school enrollment in primary education attained in recent years, additional school transfers could be, for example, exclusively oriented to high school students, accounting for the timing of program exposure and possibly exploring variations in conditionalities to extend impacts beyond high school and in other outcome variables of interest. Secondary school completion is the access gate to college education and thereby to achieve further higher returns from education.

One limitation of the study is the lack of data at the intra-household level that would permit us to identify the specific channels or mechanisms through which the received cash transfers operate within the household (among members), and how these contribute to high school completion. Similarly, a still unsettled question regarding CCTs refers to the role of conditionalities on program impact. In the case of PROSOLI, anecdotical evidence suggests program administrators have strengthened the validation process of school enrollment only in recent years; however, validation of school assistance still represents a major challenge. Meanwhile, the Ministry of Education has been working on the deployment and upgrade of the Education Administration System of the Dominican Republic (SIGERD, for its Spanish acronym), an integrated platform that centralizes schools' administrative data. Assuming a closer verification of household responsibilities positively impacts school enrollment and assistance, our estimated effects of school transfers on school completion could be further enhanced.

Lastly, one interesting result that should be explored in future work is the larger impacts on high school graduation among students attending public schools, where the transfers could be more salient compared to private schools where students have generally access to more resources both at school and home. Likewise, while we observe positive impacts of BEEP transfers on high school completion both in urban and rural areas, the apparent opposite effects of these transfers on school performance between areas, particularly on grades, should be examined in more detail as more data become available. A final natural extension of the study, conditional also on data availability, is evaluating program effects on technical education programs and college enrollment, and on labor market outcomes.

References

Abadie, A., Drukker, D., Herr, J.L. and Imbens, G.W., 2004. Implementing Matching Estimators for Average Treatment Effects in Stata. *Stata Journal* 4(3): 290-311.

Abadie, A. and Imbens, G.W., 2011. Bias-Corrected Matching Estimators for Average Treatment Effects. *Journal of Business & Economic Statistics* 29(1): 1-11.

Angrist, J. and Hahn, J., 2004. When to Control for Covariates? Panel Asymptotics for Estimates of Treatment Effects. *The Review of Economics and Statistics* 86(1): 58-72.

Angrist, J. and Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, January.

Araujo, M.C., Bosch, M. and Schady, N., 2019. Can Cash Transfers Help Households Escape an Intergenerational Poverty Trap? In Barrett, C.B., Carter, M.R. and Chavas, J.P. (Eds), The Economics of Poverty Traps, Chapter 10, 357-382. National Bureau of Economic Research.

Baird, S., Ferreira, F.H.G., Ozler, B. and Woolcock, M., 2014. Conditional, Unconditional and Everything in Between: A Systematic Review of the Effects of Cash Transfer Programmes on Schooling Outcomes. *Journal of Development Effectiveness* 6(1): 1-43.

Bang H. and Robins J.M., 2005. Doubly Robust Estimation in Missing Data and Causal Inference Models. *Biometrics* 61(4): 962-973.

Barham, T., Macours, K. and Maluccio, J.A., 2017. Are Conditional Cash Transfers Fulfilling Their Promise? Schooling, Learning, and Earnings after 10 Years. CEPR Discussion Paper No. 11937.

Barrera-Osorio, F., Linden, L.L. and Saavedra, J., 2019. Medium- and Long-term Educational Consequences of Alternative Conditional Cash Transfer Designs: Experimental Evidence from Colombia. *American Economic Journal: Applied Economics* 11(3): 54-91.

Behrman, J.R., Parker, S.W and Todd, P.E., 2005. Long-Term Impacts of the Oportunidades Conditional Cash Transfer Program on Rural Youth in Mexico. Ibero-America Institute for Economic Research, Discussion Paper No. 122, Göttingen, Germany.

Behrman, J.R., Parker, S.W. and Todd, P.E., 2011. Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A five-year follow-up of PROGRESA/Oportunidades. *Journal of Human Resources* 46(1): 93-122.

Black, S.E., Devereux, P.J. and Salvanes, K.G., 2005. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics* 120(2): 669-700.

Booth, A.L. and Kee, H.J., 2009. Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment. *Journal of Population Economics* 22: 367-397.

Breton, M., 2012. The Impact of a Conditional Cash Transfers Program on Consumption and the Food Share in the Dominican Republic. Tesis no Publicada. Universidad Carlos III de Madrid.

Cahyadi, N., Hanna, R., Olken, B.A., Adi Prima, R., Satriawan, E. and Syamsulhakim, E., 2020. Cumulative Impacts of Conditional Cash Transfer Programs: Experimental Evidence from Indonesia. *American Economic Journal: Economic Policy* 12(4): 88-110.

Canavire, G. and Vásquez, H., 2012. Impacto del Programa Solidaridad en el Mercado Laboral de la Republica Dominicana. Mimeo. Banco Central de la Republica Dominicana [in Spanish].

Fiszbein, A. and Schady, N., 2009. Conditional Cash Transfers: Reducing Present and Future Poverty. Washington, DC: World Bank.

García, S. and Saavedra, J., 2017. Educational Impacts and Cost-Effectiveness of Conditional Cash Transfer Programs in Developing Countries: A Meta-Analysis. National Bureau of Economic Research, Working Paper No. 23594.

Ibarraran, P., Medellin, N., Regalia, F. and Stampini, M., 2017. How Conditional Cash Transfers Work: Good Practices after 20 Years of Implementation. Inter-American Development Bank, Washington DC.

Imbens, G. and Rubin, D.R., 2015. Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction. Cambridge University Press, New York.

Kugler, A. D. and Rojas, I., 2018. Do CCTs Improve Employment and Earnings in the Very Long-Term? Evidence from Mexico. National Bureau of Economic Research, Working Paper No. 24248.

Lozano, M., 2012. Los Determinantes de la Deserción y Repitencia Escolar de los Hogares en Condiciones de Pobreza: Evidencia de la Encuesta de Evaluación de la Protección Social 2010. Tesis no Publicada. Universidad Católica de Santo Domingo, República Dominicana [in Spanish].

Maralani, V., 2008. The Changing Relationship between Family Size and Educational Attainment over the Course of Socioeconomic Development: Evidence from Indonesia. *Demography* 45: 693-717.

McCrary, J., 2008. Manipulation of the Running Variable in the Regression discontinuity Design: A Density Test. *Journal of Econometrics* 142(2): 698-714.

Ministerio de Economia, Planificación y Desarrollo de la República Dominicana, 2019. Boletín de Estadísticas Oficiales de Pobreza Monetaria. Año 4, No. 6 (Mayo), ISSN 2415-0312 [in Spanish].

Parker, S.W. and Todd, P.E., 2017. Conditional Cash Transfers: The Case of Progresa / Oportunidades. *Journal of Economic Literature* 55(3): 866-915.

Parker, S.W. and Vogl, T., 2018. Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico. National Bureau of Economic Research, Working Paper No. 24303.

Progresando con Solidaridad – PROSOLI, 2017. Manual Operativo Programa de Transferencias Condicionadas Progresando con Solidaridad, República Dominicana [in Spanish].

Reyes, C.R., 2014. Evaluación de Impacto del Programa ILAE sobre la Repitencia, la Deserción y la Asistencia Escolar. Tesis no Publicada. Universidad Católica de Santo Domingo, República Dominicana [in Spanish].

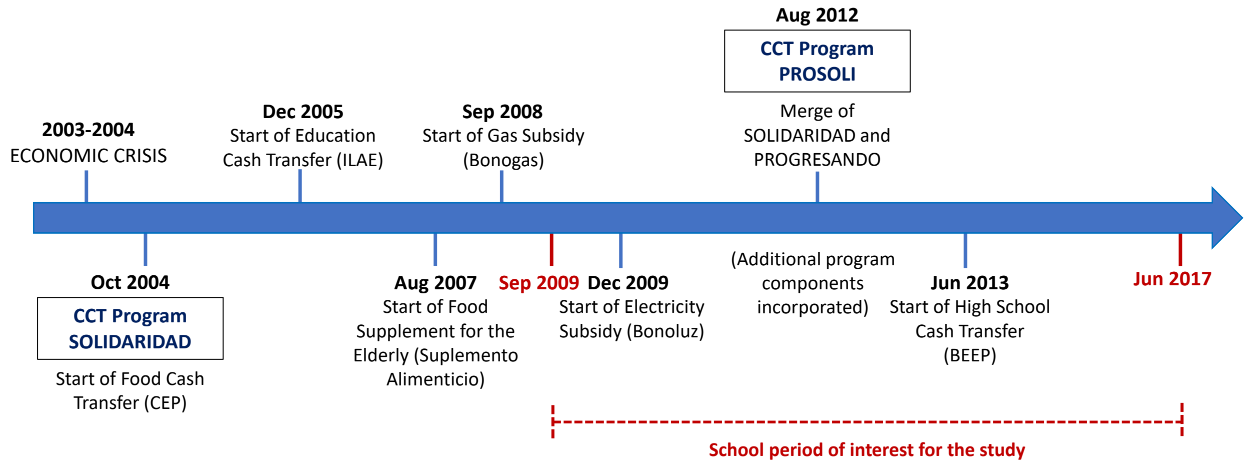
Robles, M., Rubio, M.G. and Stampini, M., 2015. Have Cash Transfers Succeeded in Reaching the Poor in Latin America and the Caribbean, Inter-American Development Bank Policy Brief No. 246, Washington DC.

Rosenbaum, P. and Rubin, D., 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1): 41-55.

Schady, N. and Araujo, M.C., 2008. Cash Transfers, Conditions, and School Enrollment in Ecuador. *Economia* 8(2): 43-70.

Vicepresidencia de la República Dominicana, 2015. Apoyo al Sistema Único de Beneficiarios para la Mejora de Levantamiento, Gestión y Análisis del Padrón de Elegibles. Reporte, Enero [in Spanish].

Figure 1. Program Timeline



Note: The additional program components incorporated in 2012 include: digital registration of family members, human formation and public awareness, environmental protection, and access to new information technologies, among others. The school period of interest for the study ranges from September 2009 through June 2017 as our working sample comprises students that finished elementary school between 2009-2010 and 2012-2013 school years and graduated or not from high school up to the 2016-2017 school year.

Table 1. Payment Scheme of PROSOLI's School Transfers

A. ILAE

Households with:	**Bi-monthly transfer per HH RD$ (US$)**
One and two students	600 (13.4)
Three students	900 (20.1)
Four or more students	1,200 (26.8)

B. BEEP

High School Format	**Grades**	**Bi-monthly transfer per student RD$ (US$)**
Regular	First and second	1,000 (22.3)
	Third and fourth	1,400 (31.3)
Technical	First to fourth	2,000 (44.7)

Source: PROSOLI Operations Manual (2017). The exchange rate RD$/US$ corresponds to the average rate for the period 2013-2017.

Table 2. Distribution of Students in Working Sample According to Transfers Bundles

Transfers Group	Freq.	%
Non-recipients	40,918	29.9
Non-conditional recipients	15,252	11.1
CEP+	17,206	12.6
ILAE+	17,970	13.1
ILAE-BEEP+	45,727	33.4
Total	137,073	100

Note: The number of observations corresponds to students that satisfy the matching cutoff to be included in the main estimation sample and whose households are found at baseline, by transfers group. The + sign indicates that households could also have received unconditonal transfers (e.g., Bonogas, Bonoluz, and Food Supplement for the Elderly). The ILAE+ and ILAE-BEEP+ group could similarly receive CEP transfers. Non-conditional recipients are elegible for unconditional transfers.

Table 3. Summary Statistics of Selected Variables in Working Sample by Transfers Group and Area of Residence

Panel A. Urban Areas

Variable	Non-recipients		Non-conditional recipients		CEP+		ILAE+		ILAE-BEEP+	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	18.10	1.78	18.28	1.80	18.51	1.87	18.71	1.97	18.27	1.76
Sex	0.49	0.50	0.47	0.50	0.49	0.50	0.46	0.50	0.51	0.50
Primary test scores	0.144	0.983	0.068	0.968	-0.011	0.96	-0.045	0.998	0.003	0.932
ICV	68.40	9.93	64.54	6.03	58.07	8.93	52.46	8.70	53.91	8.07
No toilet/latrine	0.01	0.09	0.01	0.10	0.03	0.17	0.06	0.24	0.05	0.22
HH members	4.45	1.48	4.73	1.54	4.67	1.73	5.13	1.74	5.17	1.64
School years HH members	9.99	3.35	8.67	3.10	7.45	3.31	6.27	3.11	7.03	3.10
HH head sex	0.76	0.43	0.71	0.46	0.75	0.44	0.76	0.42	0.80	0.40
HH head age	36.03	7.69	35.90	8.00	36.92	8.84	35.20	8.98	34.55	7.69
School years HH head	10.46	3.92	8.98	3.79	7.42	4.00	6.07	3.79	6.96	3.78
Presence of spouse	0.67	0.47	0.73	0.44	0.63	0.48	0.67	0.47	0.72	0.45
Years as beneficiary	0.0	0.0	8.94	0.64	6.94	3.57	10.02	2.34	10.26	1.89
Total amount transferred (US$ 2017)	0.0	0.0	1,131	1,403	1,603	1,125	3,369	1,140	4,171	1,070
High school completion	0.67	0.47	0.65	0.48	0.60	0.49	0.52	0.50	0.66	0.47
Observations	25,735		9,513		9,214		10,799		28,478	

Panel B. Rural Areas

Variable	Non-recipients		Non-conditional recipients		CEP+		ILAE+		ILAE_BEEP+	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	18.17	1.82	18.37	1.85	18.68	1.98	19.11	2.16	18.48	1.88
Sex	0.49	0.50	0.50	0.50	0.49	0.50	0.45	0.50	0.51	0.50
Primary test scores	0.116	0.957	0.069	0.958	0.042	0.986	-0.015	1.019	0.064	0.975
ICV	64.53	9.80	58.43	5.01	50.95	10.49	43.44	11.31	45.78	9.94
No toilet/latrine	0.02	0.15	0.04	0.18	0.09	0.29	0.15	0.36	0.12	0.33
HH members	4.77	1.47	4.97	1.65	5.01	1.80	5.36	1.91	5.39	1.76
School years HH members	8.56	3.37	7.32	2.96	6.03	3.22	4.73	3.03	5.60	3.04
HH head sex	0.80	0.40	0.73	0.44	0.67	0.47	0.63	0.48	0.68	0.47
HH head age	35.50	7.80	34.97	7.97	37.60	9.54	37.96	10.31	36.20	8.98
School years HH head	8.86	4.09	7.53	3.67	5.87	3.89	4.29	3.60	5.33	3.69
Presence of spouse	0.73	0.44	0.78	0.41	0.72	0.45	0.76	0.43	0.81	0.39
Years as beneficiary	0.0	0.0	8.93	0.61	7.49	3.46	10.06	2.49	10.37	1.99
Total amount transferred (US$ 2017)	0.0	0.0	1,060	1,115	1,813	1,221	3,420	1,054	4,169	958
High school completion	0.66	0.47	0.64	0.48	0.61	0.49	0.52	0.50	0.67	0.47
Observations	15,183		5,739		7,992		7,171		17,249	

Note: Summary statistics of sample variables at baseline. Average test scores are computed as the standardized difference of the mean across the four subjects examined in the first attempt of National Exams in primary school. Total amount transferred is computed as the total amount transferred during the period of analysis in RD$ at 2017 current prices and expressed in US$ using the average exchange rate for december 2017 (48.16 RD$/US$).

Table 4. Variables Used to Calculate Propensity Score

Variable Groups	Original ICV Variables
Physical characteristics of dwelling	Floor materials, wall materials, ceiling materials, dwelling type
Access to basic services	Piped water, bathroom type, electricity, garbage disposal, fuel type for cooking
Home Appliances	Fridge, tv, washer, stove
Sociodemographic characteristics	Average school years, overcrowding, share of household members under 6, share of working household members over 14, household head sex
Additional variables considered	Household head age group, school level of household head, household head literacy indicator, presence of spouse, household size, vehichle ownership, business/economic activity in dwelling, student birth year and school district fixed effects

Note: The first four group of variables correspond to those originally considered for the quality-of-life-index (ICV) computation in SIUBEN first data wave (baseline).

Table 5. Base Estimation Results of Average Treatment Effects on the Treated (ATT) on High School Attainment

Main Estimation Sample		Urban Area				Rural Area			
		Pscore matching		**Subclassification**		**Pscore matching**		**Subclassification**	
		No adjustment	*Reg adjusted*	*No adjustment*	*Reg adjusted*	*No adjustment*	*Reg adjusted*	*No adjustment*	*Reg adjusted*
A. All Students	Coeff.	0.1298	0.1174	0.1204	0.1318	0.1379	0.1244	0.1286	0.1287
	s.e.	0.0082	0.0082	0.0065	0.0071	0.0101	0.0106	0.0081	0.0082
	Obs. T=0	10,232		10,244		6,855		6,861	
	Obs. T=1	26,890		26,908		16,280		16,312	
	Avg. T = 0	0.522		0.522		0.519		0.519	
B. Male Sample	Coeff.	0.1029	0.1114	0.1088	0.1038	0.1199	0.1263	0.1281	0.1287
	s.e.	0.0112	0.0114	0.0089	0.0091	0.0135	0.0143	0.0112	0.0116
	Obs. T=0	5,465		5,487		3,717		3,748	
	Obs. T=1	13,017		13,030		7,958		7,994	
	Avg. T = 0	0.474		0.474		0.454		0.452	
C. Female Sample	Coeff.	0.1244	0.1142	0.1252	0.1243	0.0957	0.0970	0.1142	0.1088
	s.e.	0.0114	0.0118	0.0093	0.0095	0.0140	0.0141	0.0114	0.0120
	Obs. T=0	4,711		4,736		3,089		3,099	
	Obs. T=1	13,776		13,818		8,220		8,262	
	Avg. T = 0	0.578		0.579		0.602		0.601	

Note: Coefficients correspond to the pscore matching and subclassification (blocking) average treatment on the treated (ATT) estimates of exposure to additional schools transfers during high school. The dependent variable is a dummy variable that indicates if the student graduated from high school. T=1 identifies the ILAE-BEEP+ (treatment) group and T=0 the ILAE+ (control) group. The estimations include all variables used to construc the quality-of-life index (ICV) plus some additional variables: spouse presence, school level of household head, household head literacy indicator, household head age group, vehicle ownership, business (economic) activity in dwelling, student birth year and school district fixed effects; we additionally include two dummy variables that indicate whether the HH receives Bonoluz and Food Supplement for the Elderly, to control for small differences in the incidence of these two transfers between the ILAE-BEEP+ and the ILAE+ group. We implement the pscore matching estimator with no adjustment via the teffects command in Stata (caliper of 0.01 and one match only) and the adjusted pscore matching via the nnmatch comand in Stata using the bias option to adjust for differences in control variabels. In the case of the blocking estimator, for panel A, the resulting number of blocks was 17 and 11 for the urban and rural areas respectively; for panel B, 10 and 8; and for panel C, 10 and 9. For both estimators, we use the linearized pscore and we dropped observations below the first percentile of the pscore distribution for our treatment group and above the 99th percentile of the distribution for our comparison group. We use a logit specification for pscore estimation. The standard errors (s.e.) reported are heteroskedastic robust. Obs. = number of observations. Avg. = average value of dependent variable.

Table 6. Effects on High School Completion by Year of Graduation from Primary Level

Year of National Exams at Primary Level		Urban Area Pscore matching	Rural Area Pscore matching
		Reg adjusted	*Reg adjusted*
A. 2010	Coeff.	0.1162	0.1424
	s.e.	0.0164	0.0215
	Obs. T=0	2,917	1,982
	Obs. T=1	6,593	3,849
	Avg. T = 0	0.560	0.539
B. 2011	Coeff.	0.1212	0.1346
	s.e.	0.0178	0.0210
	Obs. T=0	2,329	1,673
	Obs. T=1	6,286	3,911
	Avg. T = 0	0.547	0.542
C. 2012	Coeff.	0.1126	0.1522
	s.e.	0.0175	0.0215
	Obs. T=0	2,404	1,608
	Obs. T=1	6,709	4,085
	Avg. T = 0	0.533	0.521
D. 2013	Coeff.	0.0709	0.0344*
	s.e.	0.0152	0.0204
	Obs. T=0	2,457	1,447
	Obs. T=1	6,975	4,102
	Avg. T=0	0.450	0.489
ATT Base Estimate		0.1174	0.1244

Note: Coefficients correspond to the regression-adjusted pscore matching average tretament on the treated (ATT) estimates of exposure to additional schools transfers during high school. The dependent variable is a dummy variable that indicates if the student graduated from high school. T=1 identifies the ILAE-BEEP+ (treatment) group and T=0 the ILAE+ (control) group. The standard errors (s.e.) reported are heteroskedastic robust. * indicates statistically significant at 10% level. Obs. = number of observations. Avg. = average value of dependent variable.

Table 7. Cost-Benefit Analysis ILAE-BEEP+ group versus ILAE+ group

Description	Urban area		Rural area	
	Pscore matching	Subclassification	Pscore matching	Subclassification
Estimated ILAE-BEEP+ effect on completing high school (% change)	11.7%	13.2%	12.4%	12.9%
Costs				
Average additional annual amount received by ILAE-BEEP+ group (US$ 2017)	70.41		61.94	
Benefits on employment				
Estimated change in employment probability from completing high school (source: ENCFT 2017)				
Relative to completed primary education (% change)	78.9%		60.3%	
Relative to incomplete secondary education (% change)	24.8%		18.5%	
Change in likelihood of finding a job from receiving additional US$10 per year as BEEP transfers				
Relative to completed primary education (% change)	1.3%	1.5%	1.2%	1.3%
Relative to incomplete secondary education (% change)	0.4%	0.5%	0.4%	0.4%
Benefits on salaries				
Estimated monthly salary returns from completing high school (source: ENCFT 2017)				
Relative to completed primary education (US$ 2017)	71.72		63.66	
Relative to incomplete secondary education (US$ 2017)	30.49		28.68	
Returns on monthly salary from receiving additional US$10 per year as BEEP transfers				
Relative to completed primary education (US$ 2017)	1.20	1.34	1.28	1.32
Relative to incomplete secondary education (US$ 2017)	0.51	0.57	0.58	0.60
Delayed parenthood				
Estimated delay in having first child from completing high school (source: ENDESA 2013)				
Relative to completed primary education (months)	32.88		28.24	
Relative to incomplete secondary education (months)	21.81		17.09	
Delay in having first child from receiving additional US$10 per year as BEEP transfers				
Relative to completed primary education (months)	0.55	0.62	0.57	0.59
Relative to incomplete secondary education (months)	0.36	0.41	0.34	0.36

Note: The estimated ILAE-BEEP+ effects on completing high school correspond to the base results reported in Table 5. The average additional annual amount received by ILAE-BEEP+ group, relative to ILAE+ group, is based on the total amount transferred and years as beneficiaries reported in Table 3 by group. The estimated effects on employment and salaries from completing high school are obtained from logistic and linear regression models, respectively, by urban and rural areas controlling for experience, experience squared and month fixed effects, using the 2017 Encuesta Nacional Continua de Fuerza de Trabajo (ENCFT). The estimated delay in parenthood from completing high school is obtained from a linear regression model controlling for the individual socioeconomic (poverty) condition, using the 2013 Encuesta Demografica y de Salud (ENDESA). The full set of estimations are available upon request.

Appendix A

Matching Algorithm Between School Records (PPNN) Database and Beneficiary Households (SIUBEN) Database

Due to the absence of a unique identifier between the National Exams (PPNN) and SIUBEN (second wave) databases we develop a matching algorithm based on common variables across both datasets.[48] Specifically, the algorithm uses names, last names, age, sex, and province of students in both datasets.[49] In a first stage of the matching process, the algorithm assigns a student in the PPNN database to her best match in SIUBEN database according to the score computed based on the similarity of name fields ─both first and last names─, as long as the age absolute difference between the two observations is not greater than three and both sex and province coincide. The total score is calculated as the sum of individual scores for the first and last names and it is based on Levenshtein distance. We implement this process using Python's *FuzzyWuzzy* library and its ratio function, which compares the editing distance between two string chains by computing the number of changes in terms of adding, deleting, or changing characters required to make the two chains equal. We choose this function due to its good performance when comparing short strings such as first and last names.[50] Each student is ultimately assigned to the observation in SIUBEN's with the highest score as long as the score is above the threshold embedded in the algorithm.[51]

---

[48] Given the high frequency in SIUBEN database of cases where, for instance, the first last name appeared in the cell corresponding to the second last name, the first step of the data merging process consisted in a normalization procedure for all name parts. This procedure combines into one string all parts corresponding to first and last names and then reassigns each substring to its corresponding cell considering its original position, from 1 to 4, and a frequency table of first and last names built from the national electoral roll.

[49] In the case of SIUBEN database, the province corresponds to the province of residence of the household and in the case of the National Exams database it corresponds to the school province where the student attends.

[50] For more details see https://chairnerd.seatgeek.com/fuzzywuzzy-fuzzy-string-*match*ing-in-python/ (accessed December 2021).

[51] For each observation, the algorithm organizes all its potential matches in descending order based on the score and matches it with the highest scoring candidate. In the case of a tie, the algorithm assigns the first observation found with the highest score.

The next steps explain the procedure in more detail:

1. The algorithm compares first names fields and considers whether the observation contains one or two first names in both datasets.
2. Three possible cases are evaluated:
    a. Both datasets have two first names.
    b. There exists only one first name in dataset A but two first names in dataset B.
    c. Both datasets have only one first name.
3. In the first case, both first names are compared and a minimum score of 80 is set for each pair.
4. In the second case, an initial minimum score of 80 is set, and then, the best score of the comparison between the only first name in dataset A and each of the two first names in dataset B is added to the previous minimum.
5. In the third case, an initial minimum score of 100 is set and the resulting score from the comparison of the only name in both datasets is added. The comparison of first name fields could result in a maximum score of 200 and a minimum of 160.
6. The previous steps are similarly applied to the last name fields considering the same three scenarios. Thus, paired observations could have a maximum global matching score of 400 and a minimum of 320.

After this initial matching, we performed two visual examinations to randomly selected samples. Based on these checks and with the aim of reducing the incidence of false positives' matches, we establish cutoffs above the minimum originally set by the algorithm. Therefore, we set a cutoff value of 371 for the first stage where province is binding. For the second stage, where we try to match students not paired during the first stage and province is not binding, we set a minimum score of 385.

In total, 946,992 individuals were matched between the two raw databases, which is equivalent to 53% of students in the PPNN database and 41.9% of children in SIUBEN (second wave) database. Of the matched students in the latter dataset, 219,659 completed primary school during the 2010-2013 period and satisfy the above matching criteria, which constitute our relevant sample group as BEEP high-school transfers started in the 2013-2014 school year.

Table A.1 presents summary statistics of the matching score in each stage. Of the matched students, 70% were matched during the first stage and 30% in the second stage. The mean and median of the matching score are very close within each stage, close to 385 in the first phase and 378 in the second phase. The stricter cutoff ultimately used (after the visual validations) for the first stage coincides with the 25th percentile of the calculated matching scores, while the cutoff for the second stage is more restrictive and above the median.

Table A.2 illustrates several examples of matched pairs with their corresponding matching score. The matching fields correspond to first names (first and second first names or ffn and sfn) and last names (first and second last names or fln and sln) where the first three letters in the columns headings indicate whether the variables correspond originally to the PPNN database (pnn) or SIUBEN's second data wave (sbn). The cases in Panels A and B are representative of different reference points of the *FuzzyWuzzy* score distribution in the first and second stage, respectively. Note that the mock examples in the last row in Panel A and the last two rows in Panel B are below the threshold established for each stage (371 and 385, respectively), such that these observations would have not been considered in our working sample.

Table A.1. Summary Statistics Matched Sample Between National Exams (PPNN) and SIUBEN Databases

	First Stage	**Second Stage**
Percentage of matches	70.0%	30.0%
Mean fuzzy wuzzy score	384.6	378.6
Median fuzzy wuzzy score	385	377
Standard deviation fuzzy wuzzy score	13.9	14.5
Minimum score	342	342
Maximum score	400	400
Percentile 25 score	371	370
Percentile 75 score	400	388
Threshold used	371	385

Table A.2. Examples of Matched Observations between National Exams (PPNN) and SIUBEN

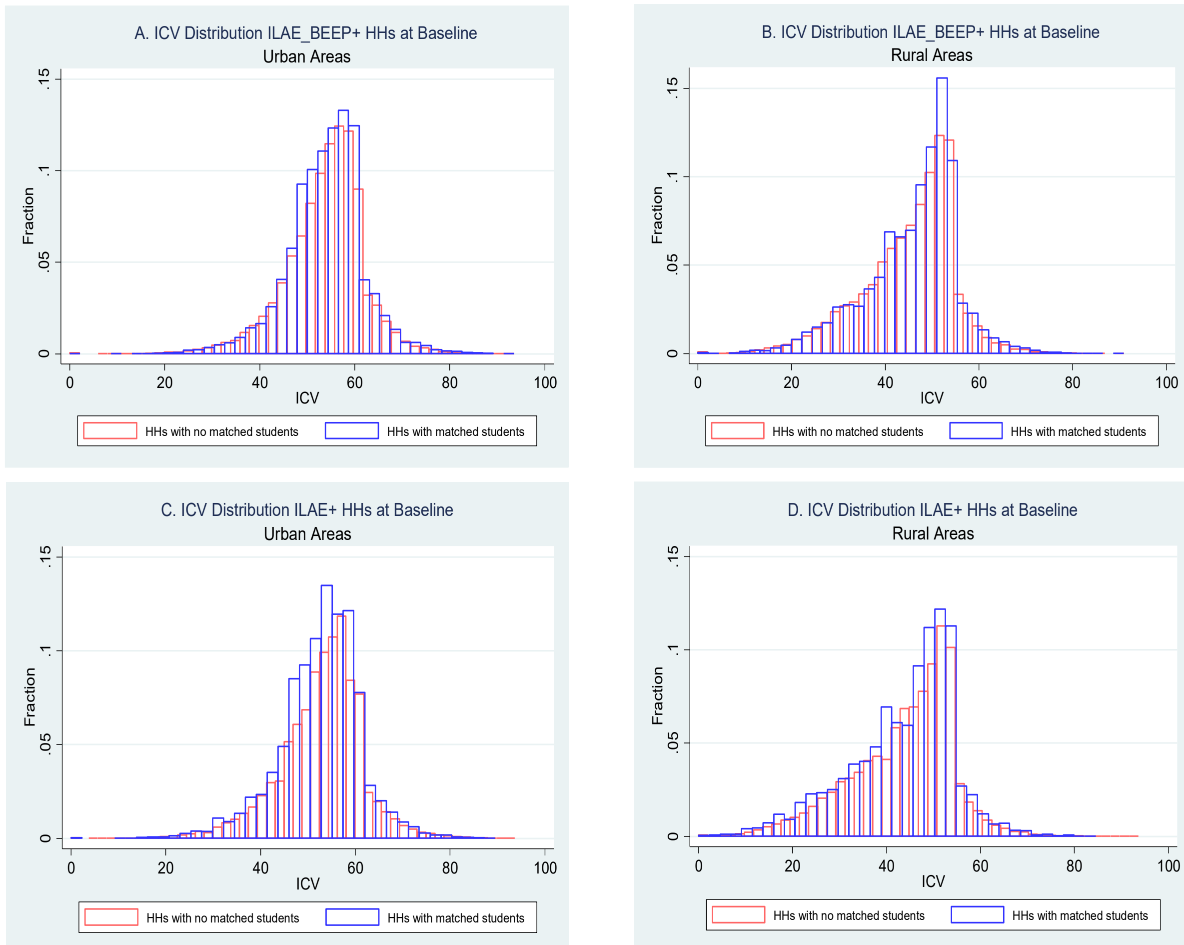
	pnn_ffn	pnn_sfn	pnn_fln	pnn_sln	sbn_ffn	sbn_sfn	sbn_fln	sbn_sln	Score
A. First stage (restricts province)	JOSEFA		VALDEZ	ROSARIO	JOSEFA		VALDEZ	ROSARIO	400
	JOSEFINA	MARIA	JIMENEZ	MENA	JOSEFINA	M	JIMENEZ	MENA	385
	YENIFER	GRISEL	MONTILLA		JENIFER		MONTILLA		371
	ANGELA	IVETTE	DIAZ	CARMONA	VIANGELA		CALMONA		342
B. Seconf Stage (does not restrict province)	RAFAELINA	ALTAGRACIA	FUENTES	LIRIANO	RAFAELINA	ALTAGRACIA	FUENTES	LIRIANO	400
	LUZ	AMANDA	BATISTA		LUZ		BATISTA		385
	MARILU		DE LOS SANTOS	BELTRE	MARILUZ		DE LOS SANTOS		377
	ESMARLIN	PAOLA	HIDALGO	PEA	ESMAILIN		PENA		342

Note: pnn = National Exams (PPNN) database; sbn = SIUBEN database; ffn = first first name; sfn = second first name; fln = first last name; sln = second last name. The actual first and last names were changed to preserve students' anonymity.

Appendix B

Supplementary Figures and Tables

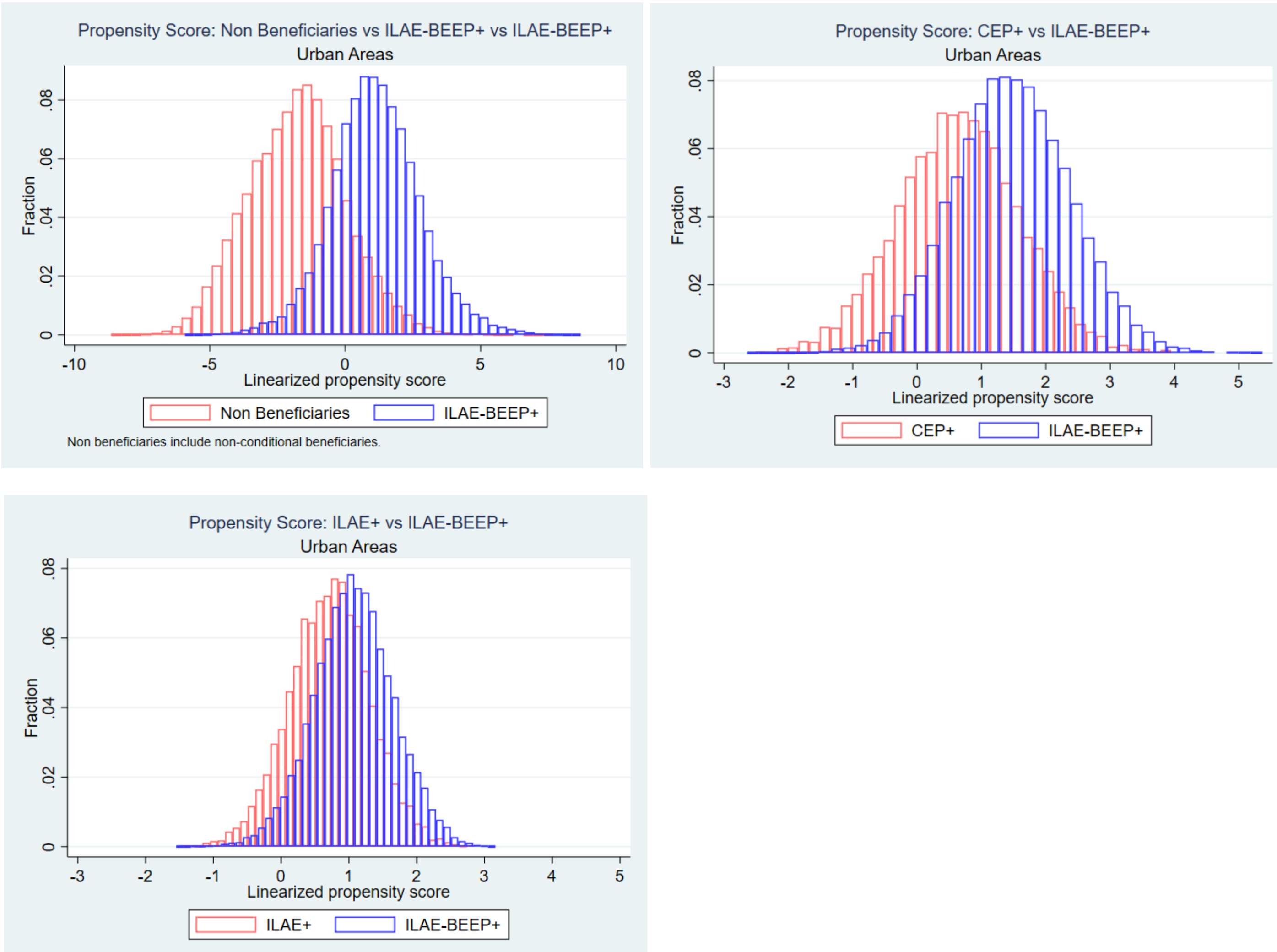
Figure B.1. Groups Comparison at Baseline: Quality-of-life Index (ICV) Distribution of ILAE-BEEP+ and ILAE+ Households with No Matched Students versus ILAE-BEEP+ and ILAE+ Households with Matched Students in Working Sample



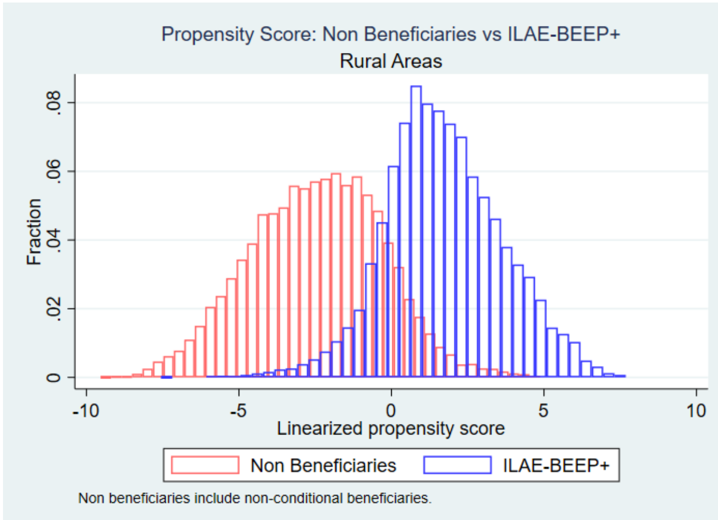
Note: ICV distributions of households (HHs) with no matched students at baseline and HHs with matched students at baseline included in our working sample for the two main beneficiary groups of interest. Figures A and B compares HHs in the ILAE-BEEP+ group in urban and rural areas respectively, and Figures C and D do the same for HHs in the ILAE+ group.

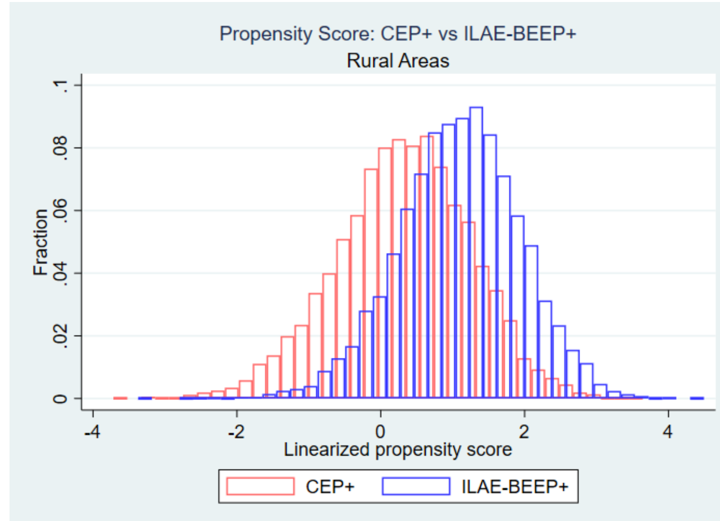
Figure B2. Propensity Score Distributions: ILAE-BEEP+ versus Other Groups

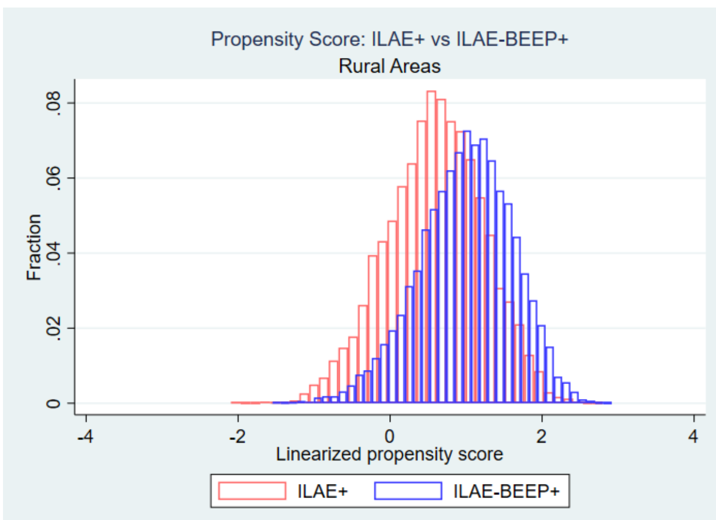
Panel A. Urban Areas



Panel B. Rural Areas







Note: Linearized propensity score distributions of households (HHs) in the ILAE-BEEP+ group against other comparison groups in urban areas (Panel A) and rural areas (Panel B). ILAE-BEEP+ HHs are compared against non-beneficiaries and non-conditional beneficiaries in Figure A, against CEP+ HHs in Figure B, and against ILAE+ HHs in Figure C.

Table B.1. Households Comparison at Baseline: Quality-of-life Index (ICV) Summary Statistics by Transfers Group at Baseline

Panel A. Urban Areas

Transfers Group	HHs with No Matched Students at Baseline	Mean (SE)	HHs with Matched Students in Working Sample	Mean (SE)
Non-recipients	435,525	65.99 (11.14)	22,456	68.47 (9.89)
Non-conditional recipients	85,317	64.39 (5.98)	8,285	64.55 (6.02)
CEP+	156,651	55.29 (9.43)	7,956	58.19 (8.93)
ILAE+	104,988	52.87 (8.73)	9,520	52.58 (8.61)
ILAE-BEEP+	56,065	53.65 (8.28)	23,287	53.96 (8.03)
Total HHs	838,546	61.36 (11.44)	71,504	60.03 (70.92)

Panel B. Rural Areas

Transfers Group	HHs with No Matched Students at Baseline	Mean (SE)	HHs with Matched Students in Working Sample	Mean (SE)
Non-recipients	243,537	57.34 (14.89)	12,981	64.61 (9.80)
Non-conditional recipients	39,830	58.47 (5.82)	4,793	58.43 (4.96)
CEP+	121,825	45.30 (11.67)	6,645	51.06 (10.52)
ILAE+	69,503	43.77 (10.81)	6,156	43.54 (11.27)
ILAE-BEEP+	35,204	45.20 (9.93)	13,743	45.95 (9.90)
Total HHs	509,899	51.87 (14.28)	44,318	53.20 (12.90)

Note: Households (HHs) in the first group corresponds to HHs in each bundle group at baseline but with no matched students from the National Exams database. The second group corresponds to HHs at baseline with matched students included in the working sample.

Table B.2. Standardized Differences in Means Before and After Matching: ILAE-BEEP+ versus ILAE+ Male Students in Base Estimation Sample

Variable	Urban Area - Men		Rural Area - Men	
	Raw	Matched	Raw	Matched
Cement floor	0.0541	0.0096	0.1019	0.0073
Granite floor	-0.0112	0.0082		
Wooden floor	-0.0009	-0.0138	-0.0093	-0.0213
Concrete wall	0.0512	-0.0103	0.0491	0.0056
Wooden wall	-0.0420	0.0173	-0.0253	0.0305
Palm Wall	0.0096	-0.0030	-0.0155	-0.0340
Concrete roof	0.0040	0.0099	-0.0080	-0.0441
Zinc roof	0.0004	-0.0090	0.0517	0.0194
Asbestos roof	-0.0248	0.0183		
Palm roof			-0.0635	0.0281
Overcrowding	-0.0614	-0.0092	-0.0601	0.0013
Tap water inside house	0.0035	0.0031	-0.0080	0.0127
Tap water outside house	0.0087	0.0062	0.0271	0.0084
Public water source	-0.0102	-0.0151	-0.0032	-0.0151
River-spring	-0.0131	-0.0048	-0.0633	-0.0124
Rain-tank-well	0.0231	0.0113	0.0581	0.0003
Water-truck	-0.0187	-0.0121	-0.0325	0.0202
Toilet	0.0452	-0.0136	-0.0270	-0.0190
Latrine	-0.0223	0.0120	0.0936	-0.0107
Power line	0.0187	-0.0246	0.0749	0.0051
Plant-generator	0.0020	0.0086	-0.0001	-0.0014
Fosil fuel generator	-0.0186	0.0207	-0.0865	-0.0090
Garbage-city	0.0574	-0.0116	-0.0121	0.0109
Garbage-priv. company	-0.0258	-0.0052	0.0123	0.0074
Garbage-burn	-0.0230	0.0032	0.0551	-0.0197
Garbage-littering	-0.0443	0.0164	-0.0580	0.0152
Cook - Propane	0.0550	-0.0263	0.1008	0.0043
Coal-wood	-0.0586	0.0277	-0.0969	-0.0067
Cook - electricity	0.0096	0.0174		
Don't cook			-0.0313	0.0165
Single unit house	0.0360	0.0191	0.0485	-0.0303
Apartment	0.0049	-0.0216		
Room	-0.0350	-0.0139	-0.0406	0.0295
Bunkhouse	-0.0317	-0.0147	-0.0286	0.0200

Table B.2. Continued

Variable	Urban Area - Men		Rural Area - Men	
	Raw	Matched	Raw	Matched
Stove	0.0590	-0.0215	0.1123	0.0200
Fridge	0.1131	0.0016	0.0988	-0.0254
TV	0.1239	-0.0141	0.1437	0.0038
Washer	0.1231	0.0006	0.1097	0.0000
Avg years education HH members	0.2160	-0.0095	0.2423	-0.0156
Share HH members < 6 years (0 - 0.25]	0.0337	-0.0177		
Share HH members < 6 years (> 0.25)	0.0241	-0.0118		
Share HH members working (= 0)	-0.0173	-0.0102	-0.0305	0.0060
Share HH members working (0 - 0.33]	-0.0245	-0.0149	-0.0411	-0.0353
Share HH members working (0.33 - 0.75]	0.0248	0.0075	0.0890	0.0061
Share HH members working (> 0.75)	-0.0083	0.0087	-0.0670	0.0217
HH head sex	0.0774	-0.0081		
Spouse	0.0832	-0.0005	0.1114	0.0115
HH head: incomplete primary	-0.1219	0.0076	-0.1707	-0.0007
HH head: full primary	0.0292	-0.0336	-0.0040	0.0068
HH head: incomplete secundary	0.1138	0.0138		
HH head: full secondary school	0.0566	-0.0117	0.1293	0.0118
HH head: incomplete college	0.0545	-0.0102	0.0584	0.0224
HH head: full college	0.0476	-0.0048		
HH head: knows how to read	0.1358	0.0154	0.1804	0.0027
HH head: age quartile 2	0.0806	-0.0018		
HH head: age quartile 3	0.0303	-0.0092	-0.0159	-0.0067
HH head: age quartile 4	-0.1173	0.0073	-0.1321	-0.0070
Private vehicle	0.0362	-0.0162	0.0504	-0.0023
Economic activity at home	0.0189	-0.0064		
HH members	0.0124	-0.0151	0.0146	-0.0517
Bonoluz	0.2023	0.0153	0.1537	0.0140
Food Supplement for the Elderly	-0.0023	-0.0133	-0.0375	0.0181
Total observations	**18,482**	**26,034**	**11,675**	**15,916**

Note: Differences corresponds to standardized differences in means before matching (raw) and after matching for male students in the ILAE-BEEP+ group vs the ILAE+ group at baseline. The table excludes variables (category groups) that are dropped in the propensity score estimation due to the lack of observations among the treatment and/or control group in urban or rural areas. Additionally, the share of HH members under 6 years of age and HH head sex were dropped in the original quality-of-life-index (ICV) model in the case of rural areas.

Table B.3. Robustness Checks: Schooling Features and Sample Construction

Sample		Urban Area Pscore matching	Rural Area Pscore matching
		Reg adjusted	*Reg adjusted*
A. Excluding students in ESD schools	Coeff.	0.1045	0.1161
	s.e.	0.0090	0.0114
	Obs. T=0	8,831	5,865
	Obs. T=1	20,608	11,987
	Avg. T = 0	0.444	0.441
B. Excluding students in technical high schools	Coeff.	0.1063	0.1232
	s.e.	0.0086	0.0110
	Obs. T=0	9,397	6,497
	Obs. T=1	23,167	15,077
	Avg. T = 0	0.479	0.493
C. Excluding students attending private schools	Coeff.	0.1328	0.1421
	s.e.	0.0093	0.0108
	Obs. T=0	8,973	6,345
	Obs. T=1	24,277	15,593
	Avg. T = 0	0.506	0.507
D. Excluding students who do not match in province	Coeff.	0.1166	0.1308
	s.e.	0.0092	0.0119
	Obs. T=0	8,154	4,920
	Obs. T=1	22,358	12,519
	Avg. T = 0	0.524	0.515
E. Student age difference <= 2	Coeff.	0.1342	0.1400
	s.e.	0.0092	0.0113
	Obs. T=0	8,486	5,745
	Obs. T=1	25,097	15,117
	Avg. T = 0	0.516	0.513
ATT Base Estimate		0.1174	0.1244

Note: Coefficients correspond to the regression-adjusted pscore matching average treatment on the treated (ATT) estimates of exposure to additional schools transfers during high school. The dependent variable is a dummy variable that indicates if the student graduated from high school. T=1 identifies the ILAE-BEEP+ (treatment) group and T=0 the ILAE+ (control) group. Panel A excludes students whose schools were included in the Extended School Day (ESD) program during its first years while they were still at primary school. Panel B excludes students at technical (vocational) high schools. Panel C excludes students attending private schools. Panel D excludes students who do not match in province during the matching process between the National Exams database and SIUBEN's first datawave (baseline). Panel E excludes matched students with absolute difference in age equal to 3. The standard errors (s.e.) reported are heteroskedastic robust. Obs. = number of observations. Avg. = average value of dependent variable.

Table B.4. Robustness Checks: Unconfoundedness Assumption

		Urban Area	**Rural Area**
		Reg adjusted	*Reg adjusted*
A. Pseudo Treatment: CEP+ vs. Non Beneficiaries and Non Conditioned Beneficiaries	Coeff.	-0.0090^	0.0203*
	s.e.	0.0097	0.0084
	Obs. T=0	30,077	17,281
	Obs. T=1	7,462	6,364
	Aver. T = 0	0.650	0.645
B. Subset Unconfoundedness	Coeff.	0.1246	0.1268
	s.e.	0.0085	0.0106
	Obs. T=0	10,250	6,861
	Obs. T=1	26,966	16,352
	Avg. T = 0	0.522	0.520
ATT Base Estimate		0.1174	0.1244

Note: Coefficients correspond to the regression-adjusted pscore matching average treatment on the treated (ATT) estimates of exposure to additional schools transfers during high school. The dependent variable is a dummy variable that indicates if the student graduated from high school. Panel A uses beneficiaries students in the CEP+ group as the treatment group (T=1) and non-conditioned plus non-recipient students as the control group (T=0). Panel B compares ILAE-BEEP+ group (T=1) versus ILAE+ group (T=0) as in the base results but uses a subset of the original controls set by excluding roof materials, lighting source, type of cooking fuel, and the additional demographic and socioeconomic variables that were not considered in the original quality-of-life index (ICV) model –roughly 38% of the initial pretreatment set–. The standard errors (s.e.) reported are heteroskedastic robust. ^ indicates not statistically significant. * indicates statistically significant at 10% level. Obs. = number of observations. Avg. = average value of dependent variable.

Table B.5. Effects on High School Completion using Other Comparison Groups

			Urban Area	**Rural Area**
			Reg adjusted	*Reg adjusted*
A. ILAE-BEEP+ vs CEP+	All	Coeff.	0.0555	0.0549
		s.e.	0.0107	0.0111
		Obs. T=0	7,748	6,966
		Obs. T=1	25,427	15,461
		Avg. T = 0	0.604	0.611
	Male Sample	Coeff.	0.0523	0.0600
		s.e.	0.0165	0.0179
		Obs. T=0	3,971	3,543
		Obs. T=1	12,613	7,657
		Avg. T = 0	0.546	0.543
	Female Sample	Coeff.	0.0672	0.0758
		s.e.	0.0150	0.0152
		Obs. T=0	3,772	3,325
		Obs. T=1	12,740	7,736
		Avg. T = 0	0.658	0.678
B. ILAE+ vs. CEP+	All	Coeff.	-0.0467	-0.0486
		s.e.	0.0118	0.0134
		Obs. T=0	8,248	7,232
		Obs. T=1	9,838	6,612
		Avg. T = 0	0.593	0.602
	Male Sample	Coeff.	-0.0571	-0.0503
		s.e.	0.0173	0.0180
		Obs. T=0	4,140	3,681
		Obs. T=1	5,254	3,529
		Avg. T = 0	0.537	0.535
	Female Sample	Coeff.	-0.0659	-0.0729
		s.e.	0.0167	0.0187
		Obs. T=0	4,047	3,540
		Obs. T=1	4,492	2,974
		Avg. T = 0	0.650	0.676

Note: Coefficients correspond to the regression-adjusted propensity score matching average treatment on the treated (ATT) estimates of exposure to additional schools transfers during high school. The dependent variable is a dummy variable that indicates if the student graduated from high school. Panel A estimates the ATT using students in the CEP+ group as controls (T=0) against students in our main tretament group: ILAE-BEEP+ (T=1). Panel B estimates the ATT using students in the CEP+ group as controls (T=0) against students in the ILAE+ group (T=1). The standard errors (s.e.) reported are heteroskedastic robust. Obs. = number of observations. Avg. = average value of dependent variable.

Table B.6. Effects on Grades – Heckman Model

	Coefficient	Urban areas: Main equation: Standardized score	Urban areas: Selection equation: If took national exam	Rural areas: Main equation: Standardized score	Rural areas: Selection equation: If took national exam
A. All students	If T=1	-0.0530	0.2900	0.0560	0.2990
	s.e.	0.0150	0.0151	0.0195	0.0188
	HH members		-0.0220		-0.0210
	s.e.		0.0047		0.0058
	Inverse Mills ratio		-0.2440		-0.2960
	s.e.		0.0214		0.0960
	Total obs.	39,257		24,413	
	Selected obs.	24,526		15,260	
B. Male students	If T=1	-0.0810	0.2520	0.071**	0.2840
	s.e.	0.0198	0.0208	0.032	0.0259
	HH members		-0.0170		-0.018**
	s.e.		0.0065		0.0081
	Inverse Mills ratio		-0.6990		-0.3340
	s.e.		0.0730		0.0138
	Total obs.	19,694		12,390	
	Selected obs.	11,125		6,788	
C. Female students	If T=1	-0.007^	0.3240	0.041*	0.3110
	s.e.	0.032	(0.022)	0.0212	0.0281
	HH members		-0.0300		-0.027
	s.e.		0.0070		0.0086
	Inverse Mills ratio		-0.307		-0.279
	s.e.		0.040		0.0126
	Total obs.	19,563		12,023	
	Selected obs.	13,401		8,472	

Note: Heckman model estimated by maximum likelihood. The main equation models the average (standardized) score obtained across four evaluated subjects (Spanish, Math, Social Sciences, and Natural Sciences). The selection equation models if the student took the national test at the end of secondary school. If T=1 is an indicator variable for ILAE-BEEP+ group and HH members are the number of household members. The Inverse Mills ratio is Heckman's lambda equal to the correlation coefficient of the error terms in the selection and main equation multiplied by the standard deviation of the main equation error term. The control variables used in the regressions are similar to the set of covariates used in the base estimations. The standard errors (s.e.) reported are heteroskedastic robust. *, ** indicate statistically significant at 10% and 5% levels. ^ indicates not statistically significant. obs = number of observations.

Table B.7. Effects on Graduation on Time – Probit Model with Sample Selection

	Coefficient	Urban areas		Rural areas	
		Main equation: If graduated on time	Selection equation: If completed high school	Main equation: If graduated on time	Selection equation: If completed high school
A. All students	If T=1	-0.026^	0.2920	0.076^	0.3000
	s.e.	0.0555	0.0151	0.0814	0.0196
	HH members		-0.0200		-0.0200
	s.e.		0.00500		0.0057
	Rho		-0.668^		-0.588^
	s.e.		0.5358		0.6026
	Total obs.	39,257		24,413	
	Selected obs.	24,442		15,229	
B. Male students	If T=1	-0.074^	0.2560	0.024^	0.2850
	s.e.	0.0454	0.0208	0.1113	0.0258
	HH members		-0.0180		-0.016**
	s.e.		0.007		0.0079
	Rho		-0.838*		-0.658^
	s.e.		0.5013		0.8879
	Total obs.	19,694		12,390	
	Selected obs.	11,073		6,772	
C. Female students	If T=1	-0.008^	0.3210	0.202*	0.3130
	s.e.	0.039	0.0223	0.1076	0.0281
	HH members		-0.0250		-0.0270
	s.e.		0.0071		0.0086
	Rho		-1.153^		-0.018^
	s.e.		0.7871		0.5922
	Total obs.	19,563		12,023	
	Selected obs.	13,369		8,457	

Note: Probit model with sample selection estimated by maximum likelihood. The main (Probit) equation models if the student graduated on time from high school (i.e., in four years or less). The selection equation models if the student completed high school. If T=1 is an indicator variable for ILAE-BEEP+ group and HH members are the number of household members. Rho is the corresponding correlation coeffecient between the error terms in the selection and main equation. The control variables used in the regressions are similar to the set of covariates used in the base estimations. The standard errors (s.e.) reported are heteroskedastic robust. *, ** indicate statistically significant at 10% and 5% levels. ^ indicates not statistically significant. obs = number of observations.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
www.ifpri.org

IFPRI HEADQUARTERS
1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org