Assessing the Long-Term Effects of Conditional Cash Transfers on Human Capital: Evidence from Colombia

IEG
INDEPENDENT EVALUATION GROUP

January 10, 2011

This document has a restricted distribution and may be used by recipients only in the performance of their official duties. Its contents may not otherwise be disclosed without World Bank authorization.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCT</td>
<td>Conditional cash transfer</td>
</tr>
<tr>
<td>ICFES</td>
<td>Instituto Colombiano para la Evaluación de la Educación (Colombian Institute for Educational Evaluation)</td>
</tr>
<tr>
<td>IEG</td>
<td>Independent Evaluation Group</td>
</tr>
<tr>
<td>IFS</td>
<td>Institute for Fiscal Studies</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least Squares</td>
</tr>
<tr>
<td>PSM</td>
<td>Propensity score matching</td>
</tr>
<tr>
<td>RD</td>
<td>Regression discontinuity</td>
</tr>
<tr>
<td>RDD</td>
<td>Regression discontinuity design</td>
</tr>
<tr>
<td>SIFA</td>
<td>Sistemas de Información de Familias en Acción (Information Systems of Familias en Acción)</td>
</tr>
<tr>
<td>SISBEN</td>
<td>Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (Selection System for Beneficiaries of State Subsidies)</td>
</tr>
</tbody>
</table>
Table 1. Effect of SISBEN Score on the Probability of Participation in Familias en Acción
Table 2. Single-Difference OLS and Matching Estimates of Impacts on High School Completion ........................................21
Table 3. RDD Reduced Form Estimates of Impacts on High School Completion .......................................................22
Table 4. OLS and Matching Estimates of Program Impacts on Mathematics ..........................................................23
Table 5. RDD Reduced Form Estimates of Impacts on Mathematics ........................................................................24
Table 6. Bounds on Program Effects on Test Scores (Matching Estimates) ..............................................................25
Table 7. Preprogram Summary Statistics by Treatment Status (Sample for the Matching Analysis) ..................43
Table 8. Logit Estimates of the Determinants of Participation in Familias en Acción (Models I and II) ............45
Table 9. Logit Estimates of the Determinants of Participation in Familias en Acción (Model III) .......................46
Table 10. OLS and Matching Estimates of Impacts on High School Completion ..................................................49
Table 11. RDD Reduced Form Estimates of Impacts on High School Completion(by gender and area) ..........50
Table 12. Upper Bounds on Program Effects on Test Scores by Gender and Area .............................................51
Table 13. RDD Reduced Form Estimates of Impacts on Mathematics .................................................................51
Table 14. OLS and Matching Estimates of Impacts on the High School Completion of Nonparticipating Young
        Adults (total and by gender and area) ................................................................................................................53
Table 15. RDD Reduced Form Estimates of Impacts on the High School Completion ........................................54
Table 16. Continuity Checks for Preprogram Household- and Individual-Level Variables ..............................56

Figures

Figure 1. Effects of the SISBEN Score on Participation in the Program .................................................................14
Figure 2. Impacts on High School Completion (RD Analysis) .................................................................................21
Figure 3. Impacts on Test Scores (RD Analysis) ......................................................................................................25
Figure 4. Merging Method—PSM Analysis ...........................................................................................................41
Figure 5. Merging Method—RDD ..........................................................................................................................42
Figure 6. Distribution of Propensity Scores on Prediction of Participation in Familias en Acción Used in Matching
        Estimation .....................................................................................................................................................44
Figure 7. RDD Estimates of Impacts on High School Completion (by gender and area) ..................................50
Figure 8. Continuity Checks for Household- and Individual-Level Variables ..................................................................................55
Figure 9. Distribution of the SISBEN Score ............................................................................................................57

Annexes

ANNEX A DATA MERGING PROCEDURES ........................................................................................................39
ANNEX B DESCRIPTIVE STATISTICS ..................................................................................................................43
ANNEX C MODELS OF PROGRAM PARTICIPATION ..........................................................................................44
ANNEX D METHODOLOGY FOR NONPARAMETRIC BOUNDS OF PROGRAM IMPACTS ..........................47
ANNEX E DISTRIBUTION OF PROGRAM IMPACTS ..........................................................................................49
ANNEX F ESTIMATES OF PROGRAM INDIRECT EFFECTS ..................................................................................53
ANNEX G ROBUSTNESS CHECKS ......................................................................................................................55
Acknowledgments

This impact evaluation was prepared by a core team led by Javier Baez and consisting of Adriana Camacho and Tu Chi Nguyen. The team was assisted by Humberto Martínez Beltrán, Román A. Zárate, and Román D. Zárate from the Universidad de los Andes.

We thank Jose F. Arias, Luis Carlos Corral, Diego Dorado, and Ana Gomez Rojas from the Colombian Department of National Planning; Julian Marino, Wilmer Martinez, Susana Ortiz, and Margarita Pena from ICFES; Omar Cajiao, Rita Combariza, Diego Molano, Juanita Rodriguez, and Hernando Sanchez from Acción Social; and other government agencies for their support in providing data, information, and helpful comments and suggestions. We also thank Catherine Rodriguez and Fabio Sanchez from Universidad de los Andes and Eric Bettinger from the School of Education at Stanford University for their comments and insights. We are grateful to World Bank staff who provided comments, especially Aline Coudouel, Theresa Jones, Margaret Grosh, Pablo Acosta, and Raja Bentaouet. The team is also grateful to Alejandro Gamboa and Carolina Renteria for the support they provided to this evaluation. Many thanks go to others inside and outside IEG and the World Bank who provided helpful comments, especially Manuel F. Castro.

The work was conducted under the general guidance of Cheryl Gray (Director) and Mark Sundberg (Manager). The team is grateful for the extensive and excellent advice provided by peer reviewers John Hoddinott (Senior Research Fellow, International Food Policy Research Institute) and Emmanuel Skoufias (Lead Economist, PRMPR, World Bank).
Evaluation Summary

Assessing the Long-Term Effects of Conditional Cash Transfers on Human Capital: Evidence from Colombia

Conditional cash transfers (CCTs) are programs under which poor families get a stipend conditional on certain actions, for example keeping their children in school or taking them for health checks. While there is significant evidence showing that they have positive short-term impacts on school participation, little is known about their long-term impacts on human capital. This paper helps fill this gap by investigating the final outcomes in education of children from poor households who benefited from Familias en Acción, a CCT in Colombia, using two nonexperimental techniques—matching and regression discontinuity design—with household surveys, a census of the poor, and administrative data. The analyses show that the program helps children, particularly girls and beneficiaries in rural municipalities, to accumulate more years of education. On average, participant children are 4 to 8 percentage points (equivalent to 8-16 percent) more likely than nonparticipant children to finish high school. Extrapolated to the current beneficiary population, this impact translates into roughly 200,000 additional students having high school diploma. Increased high school completion is also expected to improve the education, employment, and earning prospects of participants.

Regarding impacts on test scores, the analysis shows that program recipients who graduate from high school perform at the same level as equally poor non-recipient graduates in mathematics, Spanish, or the overall test. This result holds after restricting the analysis to program participants that would have finished high school even without the program to correct for selection bias induced by the “new enrollers.” The similarities between the results without and with this correction may signal an improvement in learning among the “new enrollers” who would not have gone to school without the program. However, this result also indicates the lack of additional positive impact on test scores for participant children already enrolled (the majority of beneficiaries). Even though CCTs are not designed to directly raise learning, it is relevant to explore supplementary policy actions within and outside the program to couple CCT’s objective of increasing years of education along with learning outcomes.

Objectives of this Evaluation

Conditional cash transfers (CCTs), programs that transfer money to poor families conditional upon specific education and health behaviors, have been increasingly used, particularly in Latin America, as a tool to reduce poverty and promote investments in human capital. Evidence shows that CCTs, aligned with their conditions, increased school enrollment and attendance. While these impacts are expected to serve as inputs to more human
capital accumulation, the evidence of such connection is scarce. In order to help fill this gap in knowledge, this report evaluates the impacts of \textit{Familias en Acción}, a CCT in Colombia, on final education outcomes that are closer to determining future productivity and earnings.

In particular, it seeks to answer four questions: (1) Does the additional use of educational inputs attributed to \textit{Familias en Acción} help participant children become more likely to complete high school? (2) In addition to the inherent positive effects in learning for the children that are brought into school due to the program (“new enrollers”), do those who finish high school and participate in the program perform better in academic tests than similar children not covered by the program? (3) What is the distribution of these long-term impacts across gender and rural and urban areas? (4) What are the indirect effects, if any, of \textit{Familias en Acción} on similar educational outcomes for nonparticipant young adults living in the same households as participant children?

The \textit{Familias en Acción} Program

\textit{Familias en Acción} was started in 2001-02 as an emergency safety nets program in response to a major economic crisis in the late 1990s. It then became a key element in the country’s poverty reduction strategy through improving children’s education and health. Similar to many other developing countries, Colombia was on its path toward almost universal primary enrollment and increased educational attainment. However, there remained gaps across its regions and socioeconomic groups. Education CCT, the largest component of the program, aims to reduce these gaps and build human capital. The program provides $7 and $14 monthly for children 7-18 years old in primary and secondary school, respectively, conditional upon 80 percent school attendance. The grants are targeted toward the poorest 20 percent of households as identified by SISBEN, a proxy-means index that ranks households’ welfare status. \textit{Familias en Acción} has been expanding over the years from geographic targeting (small rural municipalities with adequate facilities) to national coverage. By 2009, the program catered to approximately 2.8 million participating households at a cost of around 0.27 percent of GDP. The program is administered by a national coordinating agency (\textit{Acción Social}) and implemented by municipal governments, which are responsible for registration of participants, verification of compliance, and delivery of benefits.

Overview of Existing Evidence

There is a large amount of evidence showing consistent positive impacts of CCTs on school participation. However, little is known about the long-term impacts of CCTs on school attainment or learning. The few evaluations available in Colombia, Mexico, and Pakistan show that beneficiaries tend to be more likely to complete secondary school and achieve more schooling. For those children who would not have been in school without these programs, this inherently indicates an increase in learning. However, for the children already in school, the impacts on learning seem to be mixed, probably due to several factors that operate in opposite directions. On one hand, more time in school, better nutrition, and higher investment in educational inputs attributed to CCTs are expected to improve academic performance. On the other hand, these programs are often targeted to disadvantaged areas where the quality and supply of education are probably low. The increase in the demand for schooling could also cause classroom congestion and induce negative peer effects.

Short-term evaluations of \textit{Familias en Acción} (within 2 years of implementation), in particular, indicate that the program leads to higher consumption, higher spending on nutritious food, more children sent to school, more time devoted to studying, and infants growing taller while having fewer health issues. Similar to other CCTs, the program’s long-term effects are rarely assessed, and the results remain unclear. Largely motivated by this gap in knowledge, this report aims to investigate whether the positive impacts on education and health inputs documented so far also lead to increases in the amount of human capital and better learning.

Empirical Research Design

Since the program was not randomly allocated but assigned to the poorest households in specific municipalities, there are many inherent differences between the treatment and control groups. This evaluation uses two nonexperimental techniques
to create counterfactual groups that are comparable to the treatment groups.

1. **Propensity score matching (PSM)**

This approach is based on the baseline household survey collected in 2002 for the short-term impact evaluation of the program. The units of analysis include eligible families selected randomly in stratified comparable treatment and control municipalities. The sample is limited to children 18 years old or younger at baseline that could have finished high school during the period 2003-09.

Children in treatment and control areas are matched based on three different specifications of pre-treatment observable characteristics (household demographic and socioeconomic indicators, and community location and infrastructure) used to predict the probability of treatment. Various kernel techniques and bandwidths are used to define common supports that exclude poor matches between the two groups. Assuming there are no differences in unobservable variables that jointly influence program participation and outcomes, this method eliminates any selection bias in observable characteristics due to the nonrandom targeting of the program and decision to participate. The weighted average difference for each outcome is therefore the average impact of the program.

2. **Regression discontinuity design (RDD)**

This approach uses data from the monitoring and evaluation database (SIFA) of program beneficiaries from 2001 to 2009, and from the SISBEN surveys collected between 1994 and 2003 to construct the proxy-means test. These data are merged to create a universe of individuals below and above the threshold of eligibility as defined by the score of the proxy-means test. The sample is restricted to a group of treatment municipalities to make it comparable to the sample used in the matching analysis.

Although the SISBEN score does not perfectly predict treatment status, those below the threshold are significantly more (around 65-70 percent) likely to participate in the program than those above. Program estimates are therefore obtained from that variation in program participation as the ratio of the jump in the outcomes to the jump in the probability of treatment at the cutoff (an evaluation strategy known as “fuzzy” RDD).

The strategy relies on the assumption that children just below and just above the threshold are similar except for their participation in *Familias en Acción*.

Once the comparable treatment and control groups are identified through the two empirical methods, they are merged with the administrative records of ICFES, a national mandatory standardized test administered to students at the end of high school. Registration to ICFES is used as a proxy for high school completion since a large majority (more than 90 percent) of students in grade 11 take test and over 90 percent of the test takers end up finishing high school. Likewise, the test scores are a proxy for learning at the end of high school as they reflect the student performance in most major subjects.

**Long-Term Impacts of the Program**

*Familias en Acción* is found to help children accumulate more years of education as it increases the likelihood that participant children complete high school, on average, by 4 to 8 percentage points.

- Ordinary least squares (OLS), matching, and regression discontinuity (RD) analyses all show similar statistically significant results in the magnitude of 4 to 8 percentage points—equivalent to an increase of 8 to 16 percent—relative to the graduation rate of the control group (50 percent). An extrapolation of this impact to the current beneficiary population (approximately 3.8 million children) implies that additional 140,000 to 280,000 children could finish high school compared to the scenario with no program.

Regarding the impacts on test scores, program recipients who graduate from high school perform as well as similarly poor non-recipient graduates.

- Overall, the results indicate that program participants, compared to nonbeneficiary high school graduates (with otherwise identical socioeconomic attributes), did as well as on tests on mathematics, Spanish, or total scores given at the end of high school.

- There are still no clear differences in test scores between program participants and nonparticipants even after using bounding
procedures to address the selection bias introduced by students entering school due to the program.

- The similarities between the results without and with correction for selection bias might signal an improvement in learning among the “new enrollers” who would not have gone to school, hence would not have learned, without the program.

The program appears to have bigger positive effects on the high school completion of girls and children in rural areas.

- Both matching and RD models show that the increase in high school graduation rates is larger and more statistically significant for girls and children in rural areas.
- The lack of effects on test scores is more evident and consistent for children in urban areas while there is conflicting evidence for rural areas from the PSM and RD analyses.

The program is not found to have any effects on the high school completion rate of nonparticipating older children who live in beneficiary households.

- Matching and RD analyses do not reveal any conclusive evidence that ineligible sibling of participant children (those who were over 18 years old and had not finished high school) are more or less likely to complete high school relative to comparable adolescents who do not live in beneficiary households.

Robustness Analysis
Since Familias en Acción was not allocated randomly, there could be some potential issues that contaminate the identification of program impacts. For instance, if program beneficiaries and nonbeneficiaries differ in other aspects beyond program participation (such as educational background, income, or motivation), differences in school outcomes may be mistakenly attributed solely to the program. Statistical checks and investigation of the context of the analysis, however, indicate that the results do not seem to be undermined by these issues.

- Comparability of groups
The findings of this report rest on the assumption that the treatment and control groups, constructed through the matching and RD methods, are comparable. While there is no definitive test to rule out differences in unobservable variables, statistical tests show that both methods created groups that are rather similar on a rich set of baseline observable variables. These tests include checking for the relevance of additional variables (such as socioeconomic characteristics at the household and village level) in the models of participation within the matching analysis, as well as continuity checks within the RD analysis. Furthermore, there is no evidence of sorting around the threshold or manipulation of SISBEN scores (the main eligibility criteria) as there are no discrete changes in the score distribution around the threshold. Finally, there is no indication that non-random migration and crossover between treatment and control groups could threaten the internal validity of the analysis.

- Misspecification bias
The direction, and in most cases the magnitude, of the estimates of program effects on high school completion is stable across different specifications and functional forms of the matching and RD analyses.

- Contamination due to data matching
There could be a concern in the RD analysis that the higher likelihood of students in the treatment group being matched with ICFES registration is due to better data quality for the treated rather than their participation in the program. However, measures computed from the matching algorithms do not indicate any difference in the accuracy of data between treatment and control groups.

Discussion and Conclusions
This evaluation provides consistent evidence that a CCT in Colombia helps participant children, who have benefited from the program from one to nine years, finish high school. The intuition is that children, incentivized by the program to stay in school, progress through grades and accumulate more years of education. Besides, this positive impact of Familias en Acción is expected to translate
into further gains. First, with high school education, the participants may have access to better jobs and higher wages (by up to 7-11 percent, estimated in the Colombian context). Second, they become eligible for higher education (college and formal vocational schools), which could further improve their prospects of employability and earnings (average returns to college education in Colombia range from 7 to 18 percent increase in wages). Finally, in addition to other positive social and economic externalities, there are further dynamic benefits as more educated parents tend to have fewer children (for instance, having high school education is found to reduce the number of children by 27 percent compared to a reduction of only 14 percent for those with a primary education), and their children also have better school enrollment and attainment rates.

In addition, this paper shows that similarly poor high school graduates who benefit or not from the program perform equally in test scores, even after correcting for sample selection by excluding the children brought into school by the program. For this particular group of beneficiaries, these results indicate an improvement in their learning since they would not have learned as much had they not been in school. For the students already in school, however, this lack of positive effect may indicate the low effectiveness of the income transfer on their learning.

Although CCTs were not conceived as a tool for improving learning, it is critical to assess additional interventions to couple the objectives of increasing human capital with enhancing student performance. Some policy options to achieve this include changing the program design, for example, tying the benefits with conditions on academic performance, or accompanying the program with supplementary supply-side interventions aimed at improving school quality and increasing resources for low-performing students. Pilot tests, together with careful evaluations of these policy alternatives, would surely yield valuable knowledge about their efficacy.
1. Introduction

Motivation and Objectives

1.1 As part of the global effort to promote universal basic education, a number of programs, including conditional cash transfers, have been put in place with positive effects on school enrollment and attendance. Existing evidence shows that in many cases the use of educational services responds positively to interventions such as school construction, hiring of additional teachers, regular deworming of children, school feeding and take-home ration schemes, school vouchers, and unconditional and conditional cash transfers. In particular, conditional cash transfers (CCTs), programs that transfer money to poor families contingent on specific education and health behaviors, have been on the rise in recent years. Since 1997, more than 30 countries have adopted CCTs with the goal to reduce poverty and promote investments in human capital, although the programs are predominantly concentrated in Latin America. A recent review of the impact evaluation literature indicates that all 11 CCTs evaluated against school enrollment and 15 CCTs evaluated against attendance have positive effects (World Bank 2010). This, however, may not be surprising since most of these programs are conditional upon such school inputs.

1.2 CCTs’ impacts on children’s school participation are expected to motivate better educational attainment and performance, which are important determinants of their future productivity. If students stay in and progress through school, they could complete higher level of schooling. Furthermore, the transfers and conditions of CCTs may increase households’ disposable income and their spending on activities that are beneficial for students’ learning, such as food, books, and school supplies. There is also probably less pressure for them to work so they can spend more time on school-related activities. Finally, the value that the program places on education could be transferred to the families, enhancing their attitude toward the importance of schooling. Positive peer influence that the CCT beneficiaries receive as they attend classes could also encourage them to study harder and pursue higher education.

1.3 Nevertheless, while increased children’s school enrollment and time in school is an important input, it does not automatically translate into them attaining more education and improving learning. First, schools may be overcrowded due to the rise in enrollment, creating a negative influence on student performance. Second, the children who are brought into school by the transfers and conditions of the program probably have lower expected returns to school compared to those already enrolled since the former group may be less motivated, come from lower socioeconomic background and have less capacity or time devoted to school work. Another plausible reason for limited effects is that CCTs are often geographically targeted to poor areas where the teaching and school quality may be rela-
tively lower. Despite the substantial amount of work devoted to assessing the educational impacts of CCTs, little is known about their effects on these final outcomes, which indicate not only the amount of human capital (for instance, years of schooling) but also its quality as measured by academic performance.

1.4 This report therefore seeks to help fill in this knowledge gap and identify the expected but uncertain link between school participation and educational achievement through an impact evaluation of the long-term educational impacts of a conditional cash transfer program in Colombia. More specifically, this analysis investigates whether children from poor households that are covered by Familias en Acción and have different degrees of program exposure complete more years of education—measured by the probability of completing high school—and perform better in academic tests at the end of high school relative to similarly poor children that do not receive the transfer. Although these are not necessarily final outcomes since they do not reflect the ultimate educational achievement, they are close determinants of human welfare and economic growth. In addition, this report examines whether the effects of the program vary with the location (urban versus rural) and gender of the child. Finally, the evaluation explores possible indirect effects on the human capital of older children (above 18 years old) who by the rules of the program are ineligible for the transfer but reside in households with younger participant children.

1.5 Findings from the empirical analyses done for this study show that Familias en Acción has positive impacts on an important final education outcome—the likelihood that children complete high school. The magnitude varies from 4 to 8 percentage points depending on the research design and specifications. These impacts are more positive and significant for girls and children residing in rural areas. Therefore, by helping the “new enrollers,” who otherwise would not have finished school, go to and complete school, the program is also expected to improving their learning. When measuring test scores, the evaluation finds that program beneficiaries who finish high school perform at the same level as similarly poor nonrecipient graduates in mathematics, Spanish, or the overall test. Since the beneficiaries include both “marginal” children and children who would have gone to school regardless of the program, this finding probably means that the former group improves their learning due to the program given that they perform as well as other poor children already in school. However, when addressing the selection bias that arises from the introduction of the “new enrollers” in the treatment group, there is still no difference in test scores between beneficiaries and nonbeneficiaries. This result is of great importance for empirical and policy purposes since it might indicate that the program does not motivate better learning for children already in school, to whom the transfer only has an income effect.1

1.6 This report is structured around five chapters. The first chapter introduces the objectives of the report and the evaluation literature that it seeks to contribute; this chapter also briefly describes the program and the available evidence of its impacts. The second chapter provides an overview of the questions that this analysis seeks to answer as well as the empirical methodologies and data it uses. The third chapter presents the results from the empirical analyses, which include the program impacts on high school completion and academic performance of participant students, the heterogeneity of program impacts, and the indirect effects of the program on nonparticipating adolescents who live in the same households as participant children. The fourth chapter explains some robustness checks done to
ensure that the results are not subject to selection, misspecification, or data matching bias. Finally, the conclusion discusses different interpretations of the results as well as their implications for beneficiaries and policymakers.

Overview of Existing Evidence

1.7 There is a large amount of evidence showing that CCTs encourage households to increase the use of educational services. Impact evaluations of programs implemented in an array of countries including Brazil, Cambodia, Colombia, El Salvador, Honduras, Jamaica, Malawi, Mexico, Nicaragua, Pakistan, and Turkey indicate that, by and large, CCTs lead to immediate increases in school enrollment and attendance (World Bank 2010). Even though the size of the impacts varies with the design of the programs (amount of the transfers, conditionalities, target groups, timing of payments), the characteristics of the population (age, gender, school grade, socioeconomic status), and the conditions of the country (access to school, baseline enrollment), the direction of program effects is largely consistent across programs and evaluation methods. A subset of these evaluations also tracked the school progression of participant children relative to control children, relying mostly on data that span two years of program participation. The large majority of these studies show positive impacts in indicators such as grade progression, grade repetition, and dropout rates; but it is important to stress that these effects are more prominent among children in primary education and therefore say little about their actual accumulation of human capital in later stages of life.

1.8 However, the evidence is scant when it comes to the impacts of CCTs on educational attainment. Existing evidence from Colombia and Pakistan shows that CCT beneficiaries are more likely to complete secondary school by 4 to 7 percentage points (Alam and others 2010; Barrera-Osorio and others 2008). Only evaluations of Oportunidades in Mexico have measured the impact on rural adolescents who were old enough to have plausibly completed their schooling. A first study found that children with exposure to the benefits of approximately two years or more of benefits achieve about 0.2 grades of additional schooling (Behrman and others, 2005). Subsequent studies that look at the impacts on young adults with longer periods of exposure to the benefits of the program (nearly 10 years) show important increases in grades of schooling achieved by program participants and their labor insertion but no effects on the proportion of high school graduates going to college (Behrman and Parker, 2008; Freije and Rodriguez, 2008). Likewise, results from evaluations of programs in Cambodia and Honduras point to similar effects, yet they are estimated through simulation or analysis on samples of younger children who were still in school (Filmer and Schady 2009; Glewwe and Olinto 2004).

1.9 Similarly, the evidence of program effects on learning outcomes is thin and somewhat mixed, making it difficult to draw conclusions. Improvements in cognitive development attributed to CCTs have been consistently found only for children in pre-school. The existing literature does not identify a discernible effect on learning outcomes for older children when tested during the final grades of secondary school. This is probably due partly to practical and empirical difficulties in revisiting treatment and control children long after the program has been implemented. Furthermore, an evaluation of learning could be
confounded by selection problems. Evaluations that estimate program effects based on tests given to children in school may be biased as the treatment and control groups (beneficiary and nonbeneficiary children enrolled in school) are no longer comparable. The “new enroller” who attends school thanks to the program may be poorer and of lower ability than those already enrolled. Behrman and others (2000), for example, investigate the effects of Oportunidades in Mexico on the academic achievement of children in school, a sample prone to suffer from nonrandom selection. To address this, the analysis reweights the data to align the age and sex distributions of the treatment and control groups. The findings show that there is no effect on test scores after 1.5 years of exposure to the program. In contrast, evidence based on matching techniques from a scholarship program in Argentina, which operated as a CCT, shows that the program improved students’ performance as measured by school grades (Heinrich 2007). Among analyses that do not condition on school enrollment to avoid contamination due to compositional changes, the results show for the most part that CCT beneficiaries in secondary school do not do better in academic tests (Behrman and others 2005; Behrman and Parker, 2008 for Mexico; Filmer and Schady 2009 for Cambodia).4

1.10 This report aims to investigate whether the positive impacts on education and health inputs documented so far also lead to more schooling and better learning by early adulthood. This work adds to the existing literature in three ways. First, the analysis focuses on the dynamics of program impacts in the long run as it tracks different cohorts of treatment children who have been in the program from one to nine years. Most of the few studies that measured program effects on final educational outcomes did so with children who have been exposed to the treatment for no longer than two years.5 Second, taking advantage of the available data, this paper investigates the extent to which the final educational outcomes of older children that are not eligible for the program could be influenced by the participation of their siblings. Finally, by using data from the program’s information system for impact evaluation purposes, this paper highlights the opportunities for research that may arise from using monitoring and evaluation systems as these are becoming increasingly popular tools to administer CCTs and other safety net programs.

Background of the Familias en Acción Program

Motivation and Objectives of the Program

1.11 Familias en Acción was initially designed and implemented to mitigate the effect of the economic crisis in the late 1990s on the well-being of poor households, with the objective of preserving human capital formation. In the late 1990s, Colombia was hit by its worst economic downturn in 60 years: GDP shrank by 4.5 percent in 1999 alone; and the national poverty rate increased by 7.2 percentage points, largely erasing the gains made during the early 1990s (World Bank 2005 and 2008). A team was put together to address the social dimensions of the crisis from the Government of Colombia, the World Bank, and the Inter-American Development Bank (IADB). In response to the emergency needs of the poorest and most vulnerable groups in the country, three new safety net programs were developed: the Empleo en Acción (Employment in Action) workfare program, the Jovenes en Acción (Youth in Action) training and employment program, and the Familias en Acción (Families in Action) CCT program (Rawlings 2002). Familias en Acción was inspired by the CCT Progresa (now called Oportunidades) in Mexico, and consists of conditional subsidies to education,
nutrition, and health. Over the years, the program moved beyond the emergency response function and took on the longer-term objectives of raising the future productivity of beneficiary children through increased years of education and improved health and nutrition status. It is currently considered a key element in the social protection system and the strategy to reduce extreme poverty—the Social Protection Network to Overcome Extreme Poverty, known as Juntos.

1.12 The program started as a pilot in 2001 but was scaled up in the second half of 2002 in the context of great, but still insufficient, progress in improving educational coverage and attainment. Net enrollment rates for children aged 7-11 (primary school) increased from 77 to 93 percent from 1992 to 2002, and from 40 to 57 percent for those in the 12-17 age group (secondary school) (UNESCO, Institute for Statistics). The average educational attainment of people 15 years and older increased from 6.4 years in 1992 to 7.6 years in 2003 (World Bank 2008). This is still low, however, when compared with other countries such as Argentina (8.8 years), Panama (8.6 years), and Peru and Uruguay (7.6 years) (World Bank, 2005). In terms of school completion, the numbers have not changed significantly since the beginning of the 1990s. By 2003, still only 60 percent of children who started primary school finished fifth grade, while 57 percent of those who began secondary school finished ninth grade and only 35 percent completed eleventh grade, that is, high school (García and Hill 2009; World Bank 2008). Moreover, large inequalities exist across income levels and regions. For instance, in 2001, among children aged 12-15, around 5 percent of those from families in the top income quintile were not enrolled in school, as compared with 23 percent from the bottom income quintile. Similarly, while 25 percent of youth 16-17 years old from the top quintile were not enrolled in school, the rate was 48 percent among the bottom quintile (García and Hill 2009). In regional capitals, the average level of schooling by 2003 is 8.5 years, but it is only 5 years in the rest of the country (World Bank 2008).

STRUCTURE AND IMPLEMENTATION OF THE PROGRAM

1.13 The largest component of the program is education conditional cash transfers. Households with children aged 7-18 receive a monthly grant per child, conditional on the child attending at least 80 percent of school lessons. When the program started, the grant was Col$12,000 (approximately US$7) for children attending primary school (grades 1-5) and Col$24,000 (US$14) for those in secondary school (grades 6-11). This approximately equals the direct cost borne by low-income families to send their children to school. After the first expansion in 2005, the grant was increased to Col $14,000 and Col$28,000 respectively. The latest change in 2007 brought the grant to corresponding Col$15,000 and Col$30,000, merely keeping up with inflation. In some urban areas, the subsidies were increased to accommodate the higher opportunity cost of secondary education in the cities, and the subsidies for primary school students were replaced by nutritional subsidies for children 7-11 years old due to almost full enrollment in urban areas (Acción Social 2010a). Similar to Oportunidades, the transfers specifically target mothers—a mechanism designed to ensure that the money is invested in children and an incentive for empowering women within their communities.

1.14 Within each municipality, Familias en Acción uses a proxy means test to target the poorest households. In Colombia, all households are assigned to one of six levels of a welfare indicator, called SISBEN (Selection System for Beneficiaries of State Subsidies). The in-
Chapter 1
Introduction

dex is a function of a set of household characteristics and variables related to the consumption of durable goods, human capital endowments, and current income to calculate a score that indicates the household economic well-being. SISBEN level 1 households correspond to the extremely poor and are eligible for the program. The list of SISBEN 1 families with children 18 years old and younger serves as the basis for registration to the program and is consolidated from municipalities every three years. In 2006, SISBEN was upgraded to a new version. The program used the revised system to register new families during the expansion in 2007 and to recertify families already in the program (World Bank 2008).

1.15 Familias en Acción is implemented by Acción Social in the Office of the Presidency. The World Bank and the IADB finance part of the program through multiple loans and provide technical assistance to support the reform of Colombia’s social safety nets system. Municipal governments are responsible for ensuring adequate coordination with schools working with the program to ensure its successful local implementation. They prepare the list of families to receive the subsidies, taking into account the data on verification of compliance with the conditions. Verification involves beneficiary mothers obtaining certificates of compliance from schools every two months and delivering them to the municipal coordination office, which then sends that information to the regional and finally the national coordination unit. In each community, a committee of beneficiary mothers is elected to monitor program implementation. Local banks deliver the cash transfers to beneficiaries every two months. In total, the costs of the program represent about 0.27 percent of GDP in 2009 (World Bank 2008).

1.16 Familias en Acción has been expanding over the years to national coverage. The program initially was geographically targeted. Only municipalities that were not departmental capitals, with fewer than 100,000 inhabitants, and had access to facilities that allowed for the implementation of the program (basic education and health infrastructure and at least one bank) were eligible. Out of the 1,024 municipalities in Colombia, 691 qualified for the program. Within these communities, a total of 340,000 households in 622 municipalities were registered to participate (Attanasio and others 2006). In 2005, the program expanded to include displaced families and households in departmental capitals and municipalities which became able to offer the required services or with services accessible in nearby towns. By October 2006, the program included 847 municipalities, covering 582,000 households, as well as nearly 100,000 displaced families (World Bank 2008). Most recently, during 2007, the program expanded to municipalities with more than 100,000 inhabitants to include deprived urban areas. The program now covers 1,093 out of 1,120 municipalities in all departments in Colombia (Acción Social 2010b).

1.17 Despite its national coverage, the participation to the program is not universal. The program currently has approximately 2.8 million participating households, which corresponds to almost 62 percent of the target population. This is different from other CCT programs such as Oportunidades for which participation among the eligible population is nearly universal. There are a number of factors that could explain the low take up rates observed for Familias en Acción. Interviews of eligible families that do not participate in the program reveal that many of them either did not know about the program or missed the deadlines to register given the few days given to this process. Other factors argued by eligible families not covered by the program to limit participation include distance to centers
where banks and schools are available, high transportation costs (both to enroll and claim the benefits), lack of access to electricity, low radio and television ownership, and illiteracy (Acción Social 2009; Mina and others 2007).

**EXISTING EVIDENCE ON THE IMPACTS OF THE PROGRAM**

1.18 *Familias en Acción* incorporated an impact evaluation component at the onset of implementation, and its assessments demonstrated positive effects of the program on short-term outcomes such as household consumption and children’s school participation and nutrition status. Within two years after the program started, there was already an increase in school attendance and enrollment as well as participation in growth and development sessions, which was in line with the program conditions. School enrollment rates increased by 5-7 percentage points for children in secondary and around 2 percentage points for those in primary schools. The impacts are larger in rural areas, probably due to the lower baseline among this group. This was accompanied by an increase in the time spend in school and a reduction in the labor participation of younger children, especially in domestic work, of up to 10-12 percentage points (Attanasio and others 2006; Attanasio, Fitzsimmons, and Gomez 2005). With regard to the health and nutrition component, the program helped raise the DPT vaccination rate, lower diarrhea prevalence, and increase heights and birth weight among infants (Attanasio and others 2005). The transfers from the program also increased consumption by 13-15 percent, which is around 60 percent of the grant value. This means that the beneficiary households either saved part of the stipend or experienced some forgone income. Nevertheless, the share of food consumption was shown to increase and there was higher intake of nutritious food, which may signal a behavioral change toward investing in the well-being of the household (Attanasio and others 2009; Attanasio and Mesnard 2005).

1.19 As the program continued and expanded over time, there were a few efforts to measure its medium-term effects, but they mainly focused on the initial sustainability of the short-term outcomes rather than outcomes that are closer to determining the future welfare of beneficiaries. Indeed, more than 2-3 years into the program, the effects on school enrollment and child labor were sustained; and grade repetition was found to decrease. The impacts on diarrhea prevalence and nutrition status of infants also appear to remain. While it is not clear whether consumption continued to increase, households seem to maintain their improvements in food quantity and quality. Furthermore, there was a significant reduction in absolute poverty among the participants by up to 12-17 percentage points, resulting from the $7-14 monthly stipend per child (IFS and others 2006; García and Hill 2009). One evaluation attempted to go beyond the short-term impact by measuring the education gap of the children, defined as the difference between the highest level of education completed by the child and their expected level of education in accordance with age. It was found that registration in the program contributes to a reduction in the education gap of more than one-fifth of a year and an additional year of membership decreases the education gap by 7-9 percent of a year (Bhargava 2007).

1.20 Only evaluation that looks at the test scores of beneficiary children—a closer determinant of the quality of human capital and future productivity—shows mixed results. To account for the problem of selection arising from compositional changes, García and Hill (2009) examined the impact of *Familias en Acción* on school progression and academic
achievement for the students who would have been enrolled in school even without the program. Treated children were then matched to children in the control group based on the propensity to stay in school, derived from those in the control group that were enrolled both before and after program implementation. While fifth graders in the treatment group did better on math and language tests than those in the control group by 40-47 percent of a standard deviation, particularly in rural areas, program effects for ninth graders in both subjects are negative. Yet, the results of this paper are problematic for at least two reasons. First, if program effects do exist, they are probably difficult to identify in this analysis due to low statistical power. Sample sizes for the nonparametric models estimated in the paper vary between 100 and 300 observations depending on the age groups. Second, and perhaps more importantly, nationally administered tests used by the authors to infer the academic performance of children (known as “Pruebas Saber”) are only representative at the school rather than at the individual level.
2. Evaluation Focus and Empirical Approach

Evaluation Questions

2.1 This report studies the long-term effects of participation in the Familias en Acción program on the final educational outcomes of poor children in Colombia. As noted in the preceding chapter, the expansion of CCTs worldwide has been accompanied by strong evidence showing that they raise families’ investments in the education of children in the short term. Yet, much less is known about the connection between these improved “intermediate” outcomes and the final outcomes in education such as school completion and learning. The motivation of this report is to contribute to the understanding of this empirical subject by addressing four evaluation questions:

1. **Average long-term impacts on high school completion:** Does the additional use of educational inputs attributed to the program help children increase their stock of human capital (educational attainment)? This evaluation follows and contrasts high school completion rates of different cohorts of children in the treatment group that have varied program exposure (ranging from one to nine years) with the equivalent outcomes of comparable children in the control group.

2. **Average long-term impacts on learning:** In addition to the implicit positive effects in learning for the “new enrollers” that are brought into school due to the program, do those who finish high school and participate in the program perform better in academic tests than equally disadvantaged children not covered by Familias en Acción? The analysis follows the same approach used to assess the impacts of the program on school completion but focuses on academic performance in tests given at the end of high school. Due to methodological considerations, more robust program effects on test scores are also estimated to address selection issues caused by potential compositional changes.

3. **Heterogeneity of impacts:** If the program has positive impacts on long-term education outcomes, what is the distribution of these impacts across different socio-demographic groups? The study examines, for instance, the differences in the significance and magnitude of program impacts between participant girls and boys and between beneficiary children in rural and urban areas.

4. **Indirect effects:** Does the program have positive, negative, or no effects on the school completion of nonparticipant young adults living with eligible children? These young adults were more than 18 years old when the program started—and therefore ineligible for the program—but they have not finished high school and have been exposed to the program though the eligible children in their households. Due to data constraint, it is only possible to investigate the school completion and not the learning outcome for these children.
Research Design

2.2 The program was not randomly assigned as eligibility requirements were based on geographic targeting and a welfare indicator. As discussed before, only extremely poor households that have at least one child between 7 and 18 years old and fall under SISBEN level 1 are deemed eligible for Familias en Acción. Additionally, the program was initially implemented only in certain qualified municipalities. These eligibility criteria may be problematic for evaluation as they may induce different sources of bias including observable and unobservable factors that could be correlated with program eligibility and educational outcomes. For instance, small towns like those initially targeted by the program may have poorer public infrastructure, less dynamic economies, and therefore lower returns to schooling. Parents living in these places may be less willing to send their children to school. In that case, lower improvements in school attainment over time among participant children relative to nonparticipant children in bigger towns could be mistakenly attributed to the program as negative effects. Moreover, those families who sign up for the program in treatment areas may be different in many aspects from those who decide not to participate.

2.3 Given that the program was not assigned in a random fashion, the general strategy is to use children in control areas or nearly eligible for the program to construct the counterfactual of treated children. The empirical analysis employs two different nonexperimental techniques to contrast the school completion rates and test scores of eligible children in municipalities participating in Familias en Acción with comparable poor ineligible children that reside in municipalities with and without the program. (1) Building on the research design of the short-term impact evaluation, the first exercise estimates program effects based on a comparison of outcomes between various cohorts of children from treatment areas and a matched set of eligible children from comparable municipalities not covered by the program. (2) The second research design exploits the eligibility cutoff across households in participating municipalities to estimate the causal effects of the program. Households that lie just below and just above the cutoff score based on the proxy means test are assumed to be statistically indistinguishable from each other, but the first group receives the treatment and the latter does not. Program impacts are then obtained from this discontinuity in program participation. These two empirical approaches are discussed in more depth in the following section.

First Empirical Approach – Propensity Score Matching

2.4 The first research design follows children from the treatment and control areas identified during the short-term impact evaluation of the program. This evaluation was based on a nonexperimental design that compares eligible households from municipalities covered by the program with potentially eligible households (also classified as SISBEN 1) from selected comparable areas not targeted by the program.10 Three rounds of data (a baseline and two follow-up rounds) were collected to assess the early impacts of the program. The analysis uses data from that evaluation to follow different cohorts of children in treatment and control municipalities who were interviewed as part of the baseline survey undertaken in 2002.11 These children have different levels of exposure and, thus, would be important to measure how the impacts of the program vary with different years of exposure. However, due to empirical limitations, this marginal impact cannot be properly identified.12
This same limitation applies to the second empirical approach discussed with more details in the next section.

2.5 In order to make children comparable in a wide range of observable characteristics, matching methods are used to adjust for potential biases due to nonrandom targeting and selection into the program. Although careful procedures were followed in the early evaluation to select comparable control areas, the subsample of treatment and control children selected for the matching analysis appear to differ in some demographic, socioeconomic and other household and community-level characteristics. In fact, a comparison of baseline indicators along some relevant variables reveals important differences between the two groups (Table 7 in Annex B). To tackle this problem, the statistical analysis of this report uses econometric matching techniques in the comparison of outcomes between treated and control children. This means that the distribution of observable characteristics of the control group is transformed to resemble that of the treatment group. The standard underlying assumption of this method is that matching on the propensity score (that is, the estimated probability of participation in the program) eliminates any bias generated by pretreatment differences between the two groups as long as there are no differences in unobservable variables that jointly influence program participation and outcomes. The availability in the baseline survey of relevant pretreatment information to model program targeting and participation, in principle, makes the application of the matching methodology suitable for the evaluation of Familias en Acción.

2.6 Children are matched on the basis of three different specifications used to predict the probability of treatment. The first model (PSM 1) seeks to predict the treatment status of children (the propensity score) using standard individual and household pretreatment socio-economic and demographic characteristics such as family per capita consumption, age and gender of household head and child, parental education, the number of children in the household, and a dummy variable for single-headed households. The second model of probability of treatment (PSM 2) extends PSM 1 to add a number of municipality-level covariates that proxy for measures of educational supply and demand (for instance, access to schools and pupil/teacher ratio). Finally, the propensity to participate in the program is estimated with another parametric model (PSM 3) that includes all but the health variables used by Attanasio, Fitzsimons and Gomez (2005) to evaluate the short-term impacts of the program. This augmented specification includes additional household- and municipality-level variables mostly associated with the attributes of the dwellings, family structure, access to basic services, population, relevant public infrastructure, population, and geographic characteristics (a complete list of all variables and results from the models of participation are presented in Table 8 and Table 9 in Annex C). Regarding the common support in matching, the distributions of estimated propensity scores for each of the three different model specifications suggest that there is a strong overlap between the treatment and comparison groups (Figure 6 in Annex C).

2.7 To estimate the long-term impacts of the program, this analysis compares the probability of completing high school and the test scores of children from treatment municipalities with the same outcomes of matched children in control areas. The first outcome of interest is a dummy variable for whether children graduated from high school, proxied by their registration for the ICFES (Colombian Institute for Educational Evaluation)
test, a mandatory exam given to all students in grade 11 before they graduate from high school. For obvious reasons, this outcome (and the test score) is observed only once for the same child so there are no comparable data at baseline. Therefore, post-program single differences in high school completion between treated children and matched controls are calculated nonparametrically for each of the three propensity score models in the final sample. To assess the effects of the program on learning, an identical procedure is performed on children who have taken the test to calculate differences in exam scores between treated and control children. Under the assumption of no selection on unobservables, the weighted average difference for each outcome resulting from the matching procedures is expected to provide the average impact of the program. Various kernel techniques and bandwidths are used to match children and define common supports that exclude poor matches between the two groups. We then estimate the average treatment effect as follows:

\[
\Delta ATE = \frac{1}{n} \sum \left[ E(Y | D = 1, X) - E(Y_0 | D = 0, X) \right]
\]

Where Y stands for the outcome variable, D denotes treatment status (equals to one for treated and zero for the control group), and X is the set of covariates included in the models of participation.

However, estimating program impacts on learning based on tests given to children enrolled in school is problematic. Considering that CCTs, including *Familias en Acción*, have increased school enrollment, program participants are more likely than nonparticipants to take tests in school. However, the “new enrollers” who are brought into school and promoted through grades due to the program are probably different (for instance, less motivated or of lower ability) from those who would have enrolled regardless of the program. There could be also heterogeneity in the expected returns to education between the two groups of children. If this type of selection exists, the test score distributions of treated and control children in school are not comparable. To control for this possible sample selection, the analysis constructs nonparametric upper bounds of program effects on learning by symmetrically truncating the two distributions at some specific percentiles (Angrist, Bettiinger, and Kremer 2004). A more detailed explanation of this methodology is presented in Annex D. In turn, unadjusted comparisons of test scores from matching models provide lower bound estimates of program impacts, giving a sense of how the “new enrollers” would perform relative to similar children already in school.

### Second Empirical Approach – Regression Discontinuity Design

The second research design exploits variation in assignment to treatment arising from an eligibility rule of the program. Conceived in part as a poverty reduction program, *Familias en Acción* was explicitly targeted to the poorest households in Colombia. SISBEN, a proxy-means test designed to identify the most vulnerable population that qualifies for various social programs, was used to define the target population of the program. The index varies between 0 and 100 and is the result of an algorithm that weighs households’ variables associated with their socio-economic wellbeing. Households placed into SISBEN 1, namely with scores below 18 in rural areas and 36 in urban areas, were considered eligible for the first phase of the program. This discrete change offers an opportunity to estimate the causal effects of the program on education outcomes with a regression discontinuity design (RDD). The intuition behind this strategy is that children in households that lie just below
and just above the threshold are statistically comparable except for their participation in the program. As a result, any discontinuity in the conditional distribution of high school graduation rates and test scores at the cutoff could be interpreted as the effect of Familias en Acción.

2.10 **Data from the program’s monitoring and evaluation system and a poverty census are combined to assemble the sample of analysis.** Familias en Acción is administered by SIFA, a management system that compiles information on a number of operational aspects of the program including historical records of all beneficiaries. When merging it with the SISBEN scores, these databases allow identifying the universe of individuals in households that lie below the cutoff of eligibility and participate in the program between one and nine years (2001-09). Information on the rest of the population – those eligible but not participating and those with SISBEN scores above the threshold of eligibility - is obtained from the poverty census that was collected between 1994 and 2003 with the goal of constructing the proxy-means test. Both datasets are merged together with the administrative data on ICFES registration and scores to create the sample of analysis.\(^{18}\)

2.11 **Program estimates are obtained from a setting known in the literature as “fuzzy” design where there are significant but not perfect changes in the probability of receiving the treatment at a certain threshold.** The SISBEN score does not perfectly predict treatment since the probability of participation in the program does not change from zero to one at the cutoff. Yet, there is a significant discontinuity in the probability of assignment to treatment at the threshold of around 70-75 percentage points (Figure 1). This may happen for at least two reasons. First, take up rates were around 62 percent so participation is not universal among the eligible population. Second, households above the threshold could in principle lobby with local authorities to gain access to the program. However, as it will be discussed in the section on robustness checks, there is no evidence of nonrandom sorting arising from this or other gaming strategies that could influence the SISBEN score. Assuming away methodological issues, the average causal effect of this “fuzzy” design is provided by the ratio of the jump in the outcome variable at the cutoff to the jump in the probability of participation in Familias en Acción also at the cutoff. This is equivalent to an instrumental variable setting in which the average effect of the treatment is obtained from compliers – individuals whose participation is affected by the cutoff. To check the sensitivity of the results to different specifications, estimates of program impacts based on the RD strategy are computed using both parametric functional forms and nonparametric procedures. Besides, to restrict the sample close to the cutoff, optimal bandwidths for the nonparametric analysis were also estimated using cross-validation methods.\(^{19}\)
2.12 In order to formally test that the probability of treatment changes discontinuously at the cutoff point, different polynomial specifications of the following model were estimated:

\[
\Pr(D = 1|S = s) = \alpha + \delta T + g(s - s^*)
\]  

(1)

Where \( T = 1[S \leq s^*] \) is an index function that indicates whether the SISBEN score of the individual is below the eligibility threshold \( s^* \). Results from the first stage of the fuzzy design, summarized in Table 1, show that there is a large and significant jump in the treatment probability \( \delta \) at the cutoff of the assignment variable. In fact, and confirming the graphical analysis, all point estimates of \( \delta \) yield by the models vary from 0.65 to 0.71 with p-values in the 0.01 to 0.05 range.

Based on these results, the empirical model in the “fuzzy” design is demonstrated by the following equations system:

\[
Y_{int} = \gamma + \beta D_{int} + g_{int}(s - s^*) + \epsilon_{int}
\]  

(2)

\[
D_{int} = \alpha + \delta T_{int} + g_{int}(s - s^*) + \theta_{int}
\]  

(3)

where \( \epsilon \) and \( \theta \) are error terms independent of \( D \) and the score \( S \), and the subscripts \( i, m \) and \( t \) index individuals, municipalities and year. This analysis instruments the treatment dummy, \( D \), with \( T \) to identify the “intent-to-treat” effect from the reduced form model of \( Y \) on \( D \).
Table 1. Effect of SISBEN Score on the Probability of Participation in *Familias en Acción* (First Stage Estimates – “Fuzzy” RDD)

<table>
<thead>
<tr>
<th>Variable</th>
<th>First Stage: Dependent variable FA</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligibility</td>
<td></td>
<td>0.664**</td>
<td>0.677**</td>
<td>0.655**</td>
<td>0.712**</td>
<td>0.671**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.002]</td>
<td>[0.002]</td>
<td>[0.002]</td>
<td>[0.001]</td>
<td>[0.002]</td>
</tr>
<tr>
<td>R squared</td>
<td></td>
<td>0.66</td>
<td>0.66</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>833,203</td>
<td>833,203</td>
<td>833,203</td>
<td>833,203</td>
<td>833,203</td>
</tr>
<tr>
<td>Functional form</td>
<td></td>
<td>Quadratic</td>
<td>Cubic</td>
<td>Quartic</td>
<td>Cubic</td>
<td>Quartic</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed effects municipality</td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Difference above threshold</td>
<td></td>
<td>Quadratic</td>
<td>Cubic</td>
<td>Quartic</td>
<td>Linear</td>
<td>Linear</td>
</tr>
</tbody>
</table>

**Notes:** Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children (enrolled or not in school) who were 18 years old or younger when they joined the program and that, based on their school attainment at the preprogram time, could have achieved grade 11 between 2003 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Models include linear [(Si - S*)], quadratic [(Si - S*)^2], cubic [(Si - S*)^3], and quartic [(Si - S*)^4] specifications of the control function below and above the cutoff of eligibility S*.

### Data, Samples, and Outcome Variables

#### 2.13 This impact evaluation relies on four sources of data (a household survey, a census of the poor, and two databases of administrative records) to construct two samples of participant and nonparticipant children for the two respective empirical approaches.

- **Data for matching methods:** The empirical approach based on matching methods uses the baseline household survey collected for the short-term impact evaluation of *Familias en Acción* to identify the units of analysis. This survey is part of an effort to collect longitudinal data from a stratified random sample of eligible families in both treatment and matched control municipalities. The survey is a standard multi-topic household survey that includes questions on demographics, household structure, education, health, consumption, employment, anthropometry, housing, shocks, and community education and health facilities. The baseline survey was carried out between June and October 2002. Two follow-up surveys revisited the same households in 2003 and 2004.

This analysis draws only from the baseline survey which interviewed 6,722 households in 57 treatment municipalities and 4,562 households in 9 control municipalities. 20 The subsample for this analysis only includes 25,000 children who were born during 1975-94 and may have graduated from high school between 2003 and 2009. 21 For instance, a child that had completed grade 6 at baseline (either in a treated or control area) should finish high school (grade 11) by 2007 if the child progressed on schedule. In contrast, a child...
starting primary school in 2002 will not be able to finish high school before 2013. Therefore, the relevant cohorts of children to estimate the average impacts of the program are those who at baseline had 4 to 10 years of schooling and were 18 years old or younger (called “PSM data”). Given that 2003 is the first year in which the program was broadly implemented, the age and grade completed of the child at baseline determines the length of program participation. This means that the pool of treated individuals for the matching approach includes children with program exposure that varies from one to seven years. The baseline survey is also used to construct most of the preprogram covariates employed in the matching procedures.

- **Data for RDD methods:** The samples of analysis for the RDD approach are constructed with two different sources of data. The first is the monitoring and evaluation system (SIFA) created for administrative purposes at the onset of the *Familias en Acción*. The system is a longitudinal census of program beneficiaries from 2001 to present. To date there is information on approximately 2.8 million families currently participating in the program. The second source of information is the data collected through the SISBEN surveys carried out between 1994 and 2003 to construct the proxy-means test. Questions were asked on households’ demographics, structure, assets, human capital, labor force participation, income, and access to basic services. By 2003, the surveys covered over 25 million individuals. Data from SIFA and the SISBEN surveys were carefully merged using information on date of birth, name, and national identification number. The resulting dataset (called “SIFA+SISBEN data”) contains the universe of individuals under the threshold of eligibility (whether or not they participate in the program), the individuals above the cutoff who in theory deemed ineligible to participate, as well as binary and continuous indicators of program participation and length of exposure. In order to focus on a period of time that is comparable to the one used in the matching analysis, the final subsample is restricted to treatment municipalities that were covered during the first phase of expansion between 2001 and 2003—beneficiaries could have been covered for up to nine years until 2009.

- **Data for outcome variables:** The two resulting data sets (one used for the matching analysis and one used for the RDD approach) are merged with the administrative records on registration for the ICFES test. This exam is a nationally recognized and mandatory standardized test that is administered prior to graduation from high school. Over six million students registered and took the test between 2000 and 2009. This administrative database identifies test takers by name, date of birth and national identification number. This dataset is merged with the cohorts of children identified in the “PSM data” assembled to perform the matching analysis and with the “SIFA+SISBEN data” to implement the RDD strategy. In order to avoid problems of nonrandom mismatch, strict procedures were followed to merge the datasets including matching based on full name, birth date, national identification number, and a minimization of the phonetic Levenshtein distance (Annex A). The final matching rates are around 18 percent for the matching analysis and 28 percent for the RDD approach.
2.14 The long-term impacts of the program on the human capital of children are assessed on two outcome variables:

- **High school completion**: This outcome is measured through a dummy variable that identifies whether a child registered or not for the ICFES test during the period 2003-2009. Registration to the test, given just before graduation, is a good proxy for high school completion for several reasons. First, a large majority (more than 90 percent) of students in grade 11 take the test. Second, over 90 percent of the test takers end up finishing grade 11 (World Bank 1993; Angrist and others 2006). Finally, the test is a requirement for college entrance.

- **Academic achievement**: Conditional on ICFES registration, learning is measured by the actual performance of the students on the test. The exam is a standardized test that assesses the academic achievement of students in various subjects such as mathematics, language, biology, chemistry, physics, history, geography, and a foreign language chosen by the student. This paper focuses on the impacts of the program on the scores in the Mathematics and Language modules (35 questions each) and the total score of all modules (excluding foreign language). Test scores for the regression analysis are normalized by the mean and the standard deviation in each subject by year.
3. Impacts of the Program

3.1 This chapter reports the findings on the impacts of Familias en Acción on the final outcomes in education. As described above, program impacts are obtained from two different research designs and econometric methodologies. The first compares the outcomes of treated children and matched children from eligible families in control areas. The second approach takes advantage of the administrative criteria that determine eligibility to the program to apply an RDD. This chapter is divided into three sections corresponding to the three evaluative questions posed in chapter 2. The first section discusses whether the positive effects on the economic conditions of households and the increases in enrollment and attendance attributed to Familias en Acción — identified in previous impact evaluations — also translate into more school attainment and learning. The next section seeks to understand who among the beneficiaries may benefit more and presents the distribution of long-term program impacts disaggregated by the gender and location (urban or rural) of the child. Finally, the chapter presents evidence regarding the possible spillover effects of the program on the educational outcomes of older siblings who are ineligible for the program but may benefit from the participation of younger children in the household.

3.2 Evidence emerging from the empirical analysis of this report shows that:

- **Average long-term impacts on school completion:** Familias en Acción is found to help children accumulate more years of education as it increases the likelihood that participant children finish high school, on average, by 4 to 8 percentage points.

- **Average long-term impacts on academic performance:** The analysis shows that program recipients who graduate from high school, including those who were motivated by the program to enroll, perform as well as similarly poor nonrecipient graduates in test scores. There remains no difference in test scores between the two groups after restricting the treated to program participants that would have finished high school— even without the program. This may mean that the program improves learning for “new enrollers” (by bringing them to school) but not for children already in school.

- **Heterogeneity of impacts:** The positive impacts on high school graduation appear to accrue mostly to girls and are consistently positive for beneficiaries in rural areas.

- **Indirect sibling effects:** The program does not have an indirect effect (either positive or negative) on the probability of graduating from high school for nonparticipating older children who live in beneficiary households (those who were over 18 years old and had not finished high school at baseline).
Chapter 3
Program Impacts

Average Long-Term Impacts

Effects on High School Completion

3.3 In view of the expectation that CCTs like Familias en Acción can positively or negatively influence children’s school completion, establishing the direction of the net effect is mainly an empirical matter. There are two obvious positive direct effects, which will likely lead to an increase in the demand for education among the target population in the form of positive effects on school enrollment, attendance, and progression. One is an income effect arising from the cash transfer which increases the budget of the family so that they could afford keeping the children in school. The second comes from the program’s condition on regular school attendance, which introduces a substitution effect for children not in school or not attending regularly as it reduces the relative price of education. The first effect would be the only applicable to the participant families that are already sending their children to school regardless of the program. For the “new enrollers” who are motivated to enter school by Familias en Acción, however, both effects are relevant.28 If these predicted additional investments in educational services are continued over time, one might also expect an increase in high school completion. However, there are other possible mechanisms that could reverse these effects. For instance, perverse incentives could encourage families to delay children’s graduation just to prolong their participation in the program.29

3.4 Results from ordinary least squares (OLS) and matching analyses show that children covered by the program are more likely to complete high school than children in the control group as measured by the registration for the ICFES test (Table 2). These analyses track the different cohorts of children interviewed at baseline (2002) for the short-term impact evaluation of the program. A natural caveat, however, is the lack of baseline measures of the outcome variable, making difficult to verify whether any pre-program differences exist. Program impacts are therefore estimated using single differences in ICFES registration rates between treated and control children. The first two columns present OLS regression estimates for comparison, including a specification with no covariates (1). Overall, OLS results suggest that on average treated children are between 2.8 and 4.5 percentage points more likely to finish high school. In turn, impacts from the matching analysis are estimated using three different specifications of the model of participation. Results from these models also show positive and statistically significant effects of the program on high school completion. The estimates vary between 4 and 8 percentage points depending on the specification of the propensity score, but are not sensitive to the type of kernel used for the matching process.30 Impacts of similar magnitude (4 to 7 percentage points) have also been found in CCT and education fee waiver programs in Colombia and Pakistan.31 Since the mean high school completion rate of the control group for the period 2003-09 is 50 percent, the effect of Familias en Acción corresponds to an increase of 8 to 16 percent. This impact seems to be of high economic significance given the high dropout rates across the country—an average child starting primary school has 35-40 percent chance of finishing high school on time (Sanchez, 2010).
Table 2. Single-Difference OLS and Matching Estimates of Impacts on High School Completion

<table>
<thead>
<tr>
<th>Outcome</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school completion</td>
<td>0.028*</td>
<td>0.045***</td>
<td>0.0401 **</td>
<td>0.0840 **</td>
<td>0.0696 **</td>
</tr>
<tr>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.0187]</td>
<td>[0.0220]</td>
<td>[0.0214]</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>3,476</td>
<td>3,888</td>
<td>3,476</td>
<td>3,476</td>
<td>3,888</td>
</tr>
</tbody>
</table>

Notes: Bootstrapped standard errors reported in square brackets are obtained from 200 replications. Significant at 90(*), 95(**), 99(***) percent confidence.

The units of observation are children (enrolled or not in school at baseline) who were 18 or below at baseline (2002) and that, based on their school attainment at the preprogram time, could have achieved grade 11 between 2003 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Units of analysis are matched on the propensity score from three different specifications of a logistic regression on participation in the program. Preprogram covariates of each specification of the logit models of participation are listed in Table 8 and Table 9 in Annex C. OLS coefficients in models 1 and 2 are estimated from specifications without any covariates and with all the control variables included in Model 3 of the matching analysis, respectively.

3.5 Program effects on high school completion are also estimated on a different sample around the cutoff of eligibility using an RDD. Since the score does not predict program participation perfectly, the RDD implemented in this paper follows a “fuzzy” design. Furthermore, the effect estimated with this methodology is “local” as is the case in any RDD framework. In other words, the result is only applicable to the group of individuals around the threshold. For this reason, although program impacts derived from matching and RDD analyses seek to estimate the same parameter, they are not strictly comparable. To motivate the discussion, the results of the RDD analysis are first shown in a graphical presentation. Figure 2 shows the unconditional distribution of high school completion rates of children with respect to their ranking in the proxy-means test relative to the cutoff.32 Means of high school completion rates for each value of the normalized score provide suggestive evidence that Familias en Acción had a positive effect on graduation. The visual jump in the outcome at the cutoff indicates that the discontinuous change in eligibility increases the probability of finishing high school by about 3-4 percentage points.

Figure 2. Impacts on High School Completion (RD Analysis)

Notes: The X axis presents the normalized distance of each child’s proxy-means score to the cutoff that is used to classify households as SISBEN 1 and determines eligibility to the program. The Y axis presents the probability of the child completing high school.

3.6 The results from different specifications of the RDD models estimated econometrically confirm the estimates through graphical analysis.33 Except for the most parsimonious specification for which results are marginally significant at the 10 percent level, all the other models identify a positive and strongly significant impact on the high school comple-
CHAPTER 3
PROGRAM IMPACTS

tion of children that participate in the program. Overall, the preferred specifications (models 5 and 6 in Table 3, which include quartic specifications of the control function) show that beneficiary children are between 4.9 and 6.5 percentage points more likely to finish high school. Results do not change when including municipality fixed effects and variables to control for school quality (average score in the ICFES test and the total number of students in the school by year). It is worth noting that although program impact estimates derived from the matching and RDD analyses are not fully comparable for methodological reasons, the similarity in the magnitude of the average impacts obtained from both research designs and samples is remarkable.

Table 3. RDD Reduced Form Estimates of Impacts on High School Completion

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted FA</td>
<td>0.006*</td>
<td>0.014***</td>
<td>0.026***</td>
<td>0.026***</td>
<td>0.065***</td>
<td>0.049***</td>
<td>0.037***</td>
<td>0.030***</td>
</tr>
<tr>
<td>[0.003]</td>
<td>[0.003]</td>
<td>[0.004]</td>
<td>[0.004]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.005]</td>
<td></td>
</tr>
<tr>
<td>Functional form</td>
<td>Quadratic</td>
<td>Quadratic</td>
<td>Cubic</td>
<td>Cubic</td>
<td>Quartic</td>
<td>Quartic</td>
<td>Non-par.</td>
<td>Non-par.</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>816,038</td>
<td>816,038</td>
<td>816,038</td>
<td>816,038</td>
<td>816,038</td>
<td>303,098</td>
<td>303,098</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***). percent confidence. The units of observation are children (enrolled or not in school) who were 18 or below when they joined the program and that, based on their school attainment at the pre-program time, could have achieved grade 11 between 2002 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Models include linear [(Si - S*)], quadratic [(Si - S*)2], cubic [(Si - S*)3], and quartic [(Si - S*)4] specifications of the control function below and above the cutoff of eligibility S*. Optimal bandwidths for non-parametric models were computed following a cross-validation method suggested by Fan and Gijbels (1996).

EFFECTS ON TEST SCORES

3.7 Interventions that increase school enrollment and attendance are expected to improve learning outcomes, yet this effect may be partially or fully offset by other factors. In principle, students who attend school more regularly and spend more time doing school work due to Familias en Acción are supposed to have higher academic achievement than nonrecipient children who are out of school or attend less regularly.34 In the particular case of “new enrollers,” by bringing them into school, CCTs are expected to improve their learning relative to the scenario with no education. In addition to this, programs like Familias en Acción could lead to improved learning through the effect of cash and nutritional transfers, which encourage investments in learning enhancing inputs such as books, more nutritious food, and parental time. Yet, there may be other effects running in the opposite direction. For instance, the extra influx of students may increase class size and put additional pressure on existing educational resources. In addition to congestion, increased enrollment may also affect class composition and trigger negative peer effects in learning. Moreover, CCTs are often targeted to the neediest areas where school quality may be low. The possible existence of these factors yet a lack of understanding of their net effects on learning necessitate more empirical research.

3.8 It is important to note that CCTs, in general, have not been conceived and designed with explicit objectives and incentives to raise academic performance. Instead, their primary goals are to increase the demand for schooling and help families recover the
opportunity cost for sending them to school. However, since enrollment and attendance are initial steps to increase human capital, it is both empirically important and policy relevant to also examine the second-order effects of CCTs on learning outcomes. This question is even more pertinent for those families that would have sent their children to school even without the program, to whom the CCT would mostly be a pure income transfer. Furthermore, learning is an important determinant of human development, hence affecting the beneficiaries’ future income and intergenerational socioeconomic status. Therefore, the lack of effect on learning, while does not signify a failure of the program against its stated objective, should motivate discussions on options to complement this increase in demand with improvements in education quality.

Table 4. OLS and Matching Estimates of Program Impacts on Mathematics, Spanish, and Overall Test Scores (not corrected for selection)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>OLS</th>
<th>Matching estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mathematics</td>
<td>-0.401</td>
<td>-0.265</td>
</tr>
<tr>
<td></td>
<td>[0.392]</td>
<td>[0.410]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,867</td>
<td>2,044</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.398</td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td>[0.350]</td>
<td>[0.370]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,867</td>
<td>2,044</td>
</tr>
<tr>
<td>Overall test score</td>
<td>0.179</td>
<td>0.184</td>
</tr>
<tr>
<td></td>
<td>[0.226]</td>
<td>[0.239]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,867</td>
<td>2,047</td>
</tr>
</tbody>
</table>

Notes: Test scores are normalized by the mean and the standard deviation in each subject by year. The definition of the overall test score excludes results of the foreign language test chosen by the student. Bootstrapped standard errors reported in square brackets are obtained from 200 replications. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children (enrolled or not in school) who were 18 or below at baseline (2002) and that based on their school attainment at the preprogram time could have achieved grade 11 between 2003 and 2009, the number of years needed to complete high school was lower than the number of years of treatment, and registered for the ICFES test. Units of analysis are matched on the propensity score from three different specifications of a logistic regression on participation in the program. Preprogram covariates of each specification of the logit models of participation are listed in Table 8 and Table 9 in Annex C. OLS coefficients in models 1 and 2 are estimated from specifications without any covariates and with all the control variables included in Model 3 of the matching analysis, respectively.

3.9 Initial results from the matching estimation show that children participating in Familias en Acción perform as well as equally poor nonparticipant children on academic tests at the end of high school. As before, results based on single differences in test scores are presented for unmatched and matched samples (Table 4). Since in most cases the students take the test only once, it is impossible to control for the existence of preprogram differences between the treatment and control groups. With that caveat in mind, the findings using the OLS and matching methods appear very similar. In general, there is no evidence of significant program effects on mathematics, language (Spanish), or the overall score in the test. That is, treatment children graduating from high school do not do better or worse on the test than comparable nonparticipant graduates. However, selection bias can confound this analysis due to different characteristics of the “new enrollers” who would not have joined school and taken the test without the program. This likely introduces low-scorers into the treatment group. Yet, the finding on the sample that includes this group may also mean
that the program helps “new enrollers” improve their learning given that they score as high as the students already enrolled in school.

3.10 Program impact estimates on learning outcomes obtained from the preferred model specifications in the RDD also reveal that children in Familias en Acción and similarly poor nonparticipant children perform equally, except for some suggestive evidence of negative effects in Spanish. Findings from this analysis are summarized in Table 5. Results based on functional forms with second- and third-order polynomials produce negative and significant effects of the program on scores in Spanish and the overall test. Likewise, non-parametric procedures that are estimated on an optimal bandwidth close to the cutoff also yield negative and significant effects on Spanish. However, the results do not hold when the models allow for more flexible specifications on both sides of the cutoff that include quartic polynomials. However, when looking at academic performance in Mathematics, the RD analysis shows that participant children do as well as the children in the control group. Furthermore, there is no systematic evidence of differential performance between participant and nonparticipant children based on the scores in the overall test. However, caution should be taken when interpreting all these results because of the contamination arising from the possible selection of low-scorers into the treatment group.

Table 5. RDD Reduced Form Estimates of Impacts on Mathematics, Spanish, and Overall Test Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Quadratic</th>
<th>Cubic</th>
<th>Quartic</th>
<th>Non-parametric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1. Outcome: Mathematics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>-0.019</td>
<td>-0.017</td>
<td>-0.042**</td>
<td>-0.045**</td>
</tr>
<tr>
<td></td>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.021]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>2. Outcome: Spanish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>-0.042**</td>
<td>-0.035**</td>
<td>-0.046**</td>
<td>-0.049**</td>
</tr>
<tr>
<td></td>
<td>[0.017]</td>
<td>[0.017]</td>
<td>[0.021]</td>
<td>[0.021]</td>
</tr>
<tr>
<td>3. Outcome: Overall Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>-0.021*</td>
<td>-0.021*</td>
<td>-0.036**</td>
<td>-0.042***</td>
</tr>
<tr>
<td></td>
<td>[0.012]</td>
<td>[0.012]</td>
<td>[0.015]</td>
<td>[0.015]</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>206,312</td>
<td>206,312</td>
<td>206,312</td>
<td>206,312</td>
</tr>
</tbody>
</table>

Notes: Scores are normalized by the mean and the standard deviation in each subject by year. The definition of the overall test score excludes results of the foreign language test chosen by the student. Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***).
3.11 Results remain after addressing the probability of sample selection in the matching analysis, indicating no systematic difference in test scores between control and treated children already in school. Bounding procedures were employed in the matching analysis to address the concern that the latent distributions of test scores between children brought into school by the program and those previously enrolled may differ. In particular, the analysis follows Lee (2002) and Angrist and others (2006) to estimate a set of nonparametric bounds of program impacts for specific percentiles of the score distributions. This is achieved by symmetrically truncating the distributions of the treatment and control group at the percentile \(q_0\) where selection-free control children begin having positive scores, that is, \(Y(q_{q_0}) > 0\) (Annex D). This is equivalent to estimating the impacts for students who would have taken the test even if they had not participated in the program (“always takers”). Given that the procedure rests on the assumption that the selection bias is negative when treatment effects are positive, the unadjusted conditional-on-positive comparisons of
test scores provide lower bounds of program impacts. Estimates of adjusted nonparametric upper bounds on program effects for different percentiles are presented in Table 6. For the most part, results at φ₀ show insignificant results. There are, however, some positive and significant point estimates of the effects on Spanish and the overall test score (0.098 and 0.069 standard deviations, respectively). Yet, these results hold only for one of the specifications of the model of participation (PSM1). The absence of impacts on learning found in this paper mirrors the findings from evaluations of other CCTs in Mexico and Cambodia (Behrman and others 2005; Filmer and Schady 2009).

Distribution of Program Impacts

3.12 **This section turns to the distribution of program impacts across socio-economic characteristics of program participants.** The evidence discussed so far focuses on the average impacts of the program on high school completion and learning for cohorts with different levels of exposure that range from one to seven years in the PSM analysis and from one to nine years in the RDD. Yet, differences in household constraints and preferences in response to the treatment make it likely that the effects of the program vary considerably across different subgroups of beneficiaries. Evidence regarding the distribution of these effects is therefore important to understand the dynamics of the program impacts and enhance the effectiveness and efficiency of the intervention. This paper examines the distribution of program effects by gender and location (rural or urban).

3.13 **The positive impacts of Familias en Acción on high school completion from the matching analysis are mostly driven by the large and significant effects on girls.** The analysis of heterogeneity by gender uncovers clear differences in the size of impacts between girls and boys participating in the program (Table 10 in Annex E). On one hand, results for the sample of girls based on the OLS and matching estimates always yield positive and statistically significant effects on high school graduation rates. The magnitude of the effects ranges from 4-5 percentage points in the OLS framework to 5-8 percentage points in the matching analysis. For all these cases, results are strongly significant in statistical terms at the 1 percent level. On the other hand, the pattern in impact estimates for boys is less obvious. While OLS point estimates indicate that there are no program effects, two of the three point estimates obtained from matching models (PSM1 and PSM2) appear to be just marginally significant (at the 10 percent level). Unfortunately, there are not data to empirically identify the channels that may explain the differences in the size of the impact between boys and girls. A possible explanation for this may be that the opportunity costs of educating girls are relatively lower in rural settings (for instance, if households are less dependent on the labor of the girls for farming activities) so that the marginal effects of the transfer and the conditionality are relatively larger for girls.

3.14 **Similarly, the distribution of impacts on high school completion varies by the location of program beneficiaries.** Evidence from the matching models suggests that the increase in high school graduation rates attributed to the program accrue mostly to children (most likely girls based on the previous results) whose families resided in rural areas at baseline. Whereas the effects on the samples of children in rural areas are positive and strongly significant in both economic and statistical sense in all models, analogous point es-
Estimates obtained for beneficiaries in urban areas are mostly insignificant—except for a marginally significant impact from one of the matching models (PSM2). Even if the discussion is limited to this particular model, the findings show that the size of the effect in urban centers is still half of that for children in rural municipalities—about 5 percentage points compared to 10.5 percentage points (Table 10 in Annex E).

3.15 Results from the sample used in the RD analysis also show larger program effects on high school completion among girls. As shown in Table 11 in Annex E, the impacts of the program estimated in the RD framework indicate that treated boys and girls are both more likely to complete high school relative to control children. However, the values of point estimates from the preferred RDD specifications (models 5 and 6) suggest that the effects are a little larger among girls (5.6-7.5 percentage points) than boys (4.0-5.5 percentage points). Contrary to the heterogeneity by location observed in the matching analysis, the size of the effect on high school completion appears to be larger for children who reside in urban centers even though the effect is equally positive as well as statistically and economically significant in rural municipalities. Yet, it is worth mentioning once more that program impacts identified with matching and RD designs are not fully comparable for various methodological reasons discussed above.

3.16 With respect to program impacts on test scores, a disaggregation of program estimates by location shows a clear pattern of no effects for children in urban areas while there is conflicting evidence for rural areas. Selection-adjusted estimates of matching models broken down by gender provide nonparametric bounds that replicate for the most part the conclusions drawn from the average impacts (Table 12 in Annex E). That is, bounds for the subsamples of boys and girls do not reveal any differences in academic performance relative to the counterparts in the control group. A few point estimates of the upper bounds show improvements in test scores in the overall test for male beneficiaries and Spanish for female beneficiaries; yet, it is hard to conclude on the basis of these bounds because they are either too wide—while the upper bounds are positive and significant, their corresponding lower bounds are statistically insignificant—or very sensitive to the specification of the model of participation used to predict the propensity score. When looking at children in rural areas, estimates of the upper bounds suggest that those covered by Familias en Acción and that had taken the test regardless of their participation in the program did better in Mathematics, Spanish, and the overall test. However, it is difficult again to conclude since all the lower bound estimates of the impacts for this subsample are not statistically different from zero. Finally, results from the RD analysis seem to confirm the lack of impacts observed in the matching analysis for either boys or girls. Likewise, the analysis does not reveal any effects on the test scores of beneficiaries located in urban areas. In contrast, RDD estimates obtained from individuals in rural municipalities show that participant children perform worse in Spanish and the overall test when compared to those children right above the threshold of eligibility (Table 13 in Annex E). This latter result is consistent across different model specifications and subsamples.
Are there Indirect Effects on the Schooling of Nonparticipating Young Adults?

3.17 Conditional cash transfers may also have indirect effects on the schooling outcomes of nonparticipant young adults who live in households with eligible children. Theoretical models of schooling decision highlight three distinct channels that can affect the demand for schooling. The first one is an income effect that increases the liquidity of budget-constrained households and thus may have positive impacts on household consumption and investments, including investments in the human capital of eligible and ineligible children. The second is a substitution effect caused by the program rule that conditions the transfer on the regular school attendance of participant children. As a result, the conditionality increases the relative price of leisure but only for eligible children. The third effect is a displacement effect that, although positive for participant children, is expected to run in the opposite direction for ineligible children. A clear example of the latter is the reallocation of labor or child caring away from participant children—so that they are able to comply with the conditionality—towards ineligible children in the household. Therefore, given the interaction of these three different factors, the net effect of conditional transfers on the human capital of ineligible children is ambiguous. Existing empirical studies have provided evidence of either negative effects or no effects of CCTs or similar programs on the school enrollment of ineligible siblings in Cambodia, Colombia and Pakistan (Ferreira and others 2009; Barrera-Osorio and others 2008; Alam and others 2009).

3.18 This paper investigates the indirect effects of Familias en Acción on the school completion of nonparticipating young adults that reside with participant children. The analysis employs the same research designs used for eligible children. That is, the treatment group was comprised of ineligible young adults who were more than 18 years old and had not finished high school when at least one of the children of the same household joined the program. The high school graduation rate of this group is compared with those of young adults of similar characteristics that reside either in eligible families in control areas (matching analysis) or in families that were ineligible for the program but similar otherwise to participant households (RDD). The analysis draws on the same data sources used in the rest of the paper and subsamples for the econometric models are selected based on similar criteria.

3.19 Overall, the analysis does not reveal consistent evidence of either positive or negative indirect effects of the program on the school completion of nonparticipating young adults. A subset of results from the matching analysis indicates that young adults who belong to households with participant children are between 3 and 4 percentage points more likely to finish high school. However, these results do not hold for augmented specifications of the models of program participation that add mostly community characteristics to the set of control variables (Table 14 in Annex F). In turn, results from the RD approach in general show positive but statistically insignificant impacts, including results from models with high-order polynomials of the control function (Table 15 in Annex F). The only exception to this is model 5, but, and analogous to the pattern observed for PSM models, the positive effect disappears when RD models control for fixed effects at the municipality level. Likewise, there is high dispersion in point estimates when the sample is broken down by gender and area. In view of this, it is hard to conclude with an adequate level of confidence that the program has consistent positive indirect effects on the educational prospects of young
adults who do not participate in the program, live with other beneficiaries, and have not completed high school.
4. Robustness Analysis

4.1 Recognizing that this evaluation relies on nonexperimental approaches, this chapter discusses the robustness of the findings to a number of issues that may affect the internal validity of the analysis. In most cases, a randomized experiment is the optimal strategy to identify treatment effects by balancing observable and unobservable characteristics of subjects in the treatment and control groups. However, as noted above, for several reasons Familias en Acción was not randomly allocated, which may introduce spurious causality and bias. For instance, if program beneficiaries and nonbeneficiaries differ in other aspects beyond program participation (for instance, educational background, income, and motivation), changes in school attainment may be mistakenly attributed solely to the effects of the program. This paper has attempted to overcome the identification issues brought about by the nonrandom nature of assignment to treatment with two different nonexperimental research designs. This chapter assesses the validity of the causal relationship between participation and eligibility to the program and the outcome variables derived from the empirical analysis. The main conclusion is that the findings of the evaluation appear to be robust to issues such as selection, misspecification bias and contamination in data matching.

Comparability of Groups

4.2 The main identifying assumption of this paper rests on the comparability of the treatment and control groups constructed for the analysis. The first approach—matching on the propensity score (the estimated probability of assignment to treatment)—removes any bias that could arise from pre-program observable differences between the treatment and comparison groups, and the underlying assumption is that there are no unobservables that lead to nonrandom selection into the program. In other words, matching techniques only remove bias due to observable characteristics. As for the RDD approach, the analogous condition for identification is that the groups of people right below and above the threshold of eligibility need to be statistically equivalent and that the only difference between them is the treatment itself.

4.3 Even though there is no definitive test to rule out selection on unobservables, this paper carefully models program participation and conducts checks on the internal validity of the matching analysis. First, a rich set of pretreatment characteristics at the individual, household, and community levels were used to predict participation in the program. Balancing checks were also conducted to assess the comparability of treatment and control children after conditioning for observable characteristics that explain participation in the program. Regression analysis shows that, after controlling for the probability to participate, no additional conditioning variables help predict the receipt of treatment. Furthermore, there are no major statistically significant differences in the conditioning variables between the treatment and control within the same strata of similar probabilities of program participation. These balancing tests are particularly robust for the first specification of the model of
participation (PSM1). Although these tests are not definitive, they strongly support the use of matching with these propensity scores. As for the second empirical approach, impacts obtained from RD models were also estimated through specifications with several control variables including municipality fixed effects. Finally, the main findings of the paper regarding the impacts on high school completion appear to be stable across different model specifications, signaling robustness to a variety of covariates.

4.4 Various continuity checks were carried out to evaluate the identifying assumption of the RD approach, namely the comparability between households on either side of the cutoff of program eligibility. This paper has already shown that the scoring system (SISBEN) used to classify household induces a substantial jump in program participation at the threshold. Besides this difference in the eligibility, hence participation in the program, there does not seem to be any difference in other observable characteristics between the two groups on each side of the cutoff. Figure 8 in Annex G shows regression analysis on the relationship between the score and several socioeconomic characteristics available in the SISBEN survey. Because of the large sample size used in the analysis, these differences are very precisely estimated. Most of them are statistically insignificant (Table 16 in Annex G). In cases where the differences are statistically significant, the magnitude of the discontinuity is very small in an economic sense.

4.5 Another potential concern in the RDD is the possibility that individuals could manipulate the assignment variable (the SISBEN score) and generate nonrandom sorting around the threshold. An RDD approach may be invalid if more motivated and education-driven people seek to influence the value of their scores by, for instance, hiding assets, “borrowing” children from other families to increase the size of the household, or bribing local authorities and program administrators. This may undermine the comparability of people on each side of the cutoff. While it is impossible to fully rule out this type of behavior in the context of Familias en Acción, there are several reasons to believe that direct manipulation of the assignment mechanism is not a major concern. First, the score is exclusively calculated by a public agency and, with a few exceptions, the algorithm that weighs the different components of the welfare index has been kept secret. Second, an examination of the density of the score itself shows that there are no jumps in the distribution at the threshold for the whole sample or when broken down by gender and location (Figure 9 in Annex G). Third, and related to the continuity checks discussed above, there is no evidence of discrete changes in the distribution of other observable dimensions beyond the probability of participation in the program that could indicate some degree of manipulation of the scoring and ranking system.

4.6 Finally, there is no evidence that nonrandom migration, crossover, and changes in education supply may contaminate the findings of the analysis. In theory, the internal validity of the analysis might be undermined if families and/or children in control areas that are systematically different from others try to move to treatment areas to become eligible for the program. Additionally, systematic differences in school supply over time (quantity and quality of schools and teachers) that are correlated with the treatment may also confound the impacts. There are reasons, however, to think that these issues are unlikely to happen. First, migration from control to treatment municipalities is almost zero among households interviewed at baseline and resurveyed in the two subsequent rounds for the PSM ap-
proach. Second, there is no evidence showing changes in school supply that are systemati-
cally different between the treatment and control areas. Third, migration and changes in
school supply do not affect the RD analysis since the control group is comprised of children
and adolescents that are above the cutoff of eligibility but reside in the same municipalities
as participating children. Fourth, the economic incentives induced by the amount of the
transfer (around $7-14 per eligible child) are probably not enough to compensate for the di-
rect and indirect costs of migration. These observations also seem to be confirmed by pre-
vious evidence on Familias en Acción as well as a few anecdotes gathered for this evaluation.

Misspecification Bias

4.7 Problems of misspecification of the underlying regression models may introduce
bias in treatment effects, particularly in the RD design. In order to avoid such problems,
program impacts estimated in the empirical analysis of this paper were obtained from vari-
ous alternative specifications. The discussion of the most reliable findings focuses on those
that are not sensitive to either major or minor changes in the functional form. Overall, the
existence or lack of program effects on school completion and test scores, respectively, from
the first approach is fairly stable across parametric (OLS) and nonparametric (matching)
models. Similarly, and given that the consequences of model misspecification are more se-
rious in the RDD approach, the empirical models in this design were run for a number of
low-order polynomial functions as well as other more flexible functional forms including
third- and fourth-order polynomials. As shown in chapter 3, the direction—and in some
cases the magnitude—of RDD estimates is robust to the inclusion of higher order polynomi-
al terms in the control function.

Contamination Due to Data Matching

4.8 Finally, another concern is that the higher rates of matching survey data and test
records observed for the treatment group may be driven by differences in the merging
procedures and quality of information rather than by the effect of the program. Since part
of the information for the treatment and control groups comes from different datasets, it
could be that the individual-level variables used to merge the different datasets may be
more accurate for people in the treatment group. For instance, surveyors that collected the
data for the short-term impact evaluation could have been more careful to correctly keep the
information to identify individuals in the treatment group for the subsequent rounds of data
collection. Similarly, as part of the regular updates of the data entered in the information
systems of Familias en Acción (SIFA), program administrators may be more likely to correct
mistakes in names, birth dates, and national identification numbers of program benefici-
aries. As a result, individuals in the treatment group may be more likely to be matched with
records in the ICFES database, not because they are more likely to take the test, but because
of better information quality. In practice, however, measures computed from the matching
algorithms such as the Levenshtein distance do not indicate systematic differences in the
accuracy of data between the treatment and control groups. Therefore, it seems unlikely that
this is a reason to find children covered by the program to be more likely to be matched to
ICFES records.
5. Conclusions

5.1 This paper presents an evaluation of the long-term effects of *Familias en Acción*, a conventional CCT program implemented in Colombia, on the accumulation of human capital. Despite growing efforts directed to assessing the impacts of CCTs on education, gaps in knowledge exist as to whether the largely documented positive effects on enrollment and attendance are sustained over time and result in more school attainment. Equally scant is the evidence on the relationship between higher utilization of school inputs due to CCTs and similar interventions and learning outcomes. This paper seeks to help fill these gaps by empirically investigating the schooling trajectories and academic performance of various cohorts of participant children who have different levels of program exposure ranging from one to nine years. In order to do so, the impact evaluation exploits two different nonexperimental techniques (matching and RDD) to identify: (1) the average impacts of the program on high school graduation rates and test performance; (2) the heterogeneity of program impacts by gender and location of beneficiaries; and (3) the existence of indirect effects on the educational outcomes of ineligible young adults in participating households. By drawing on data from the program’s information system, this paper also highlights the opportunities for research that may arise from monitoring and evaluation systems as they become more widely used to administer safety net programs.

5.2 The quantitative analysis yields systematic evidence that *Familias en Acción* increases school attainment by helping participant children to finish high school. Indeed, treated children are on average between 4 and 8 percentage points—equivalent to 8-16 percent—more likely to graduate from high school relative to similar children in the control group. If at present the program supports nearly 3.5 million poor children and only about half of them are expected to finish high school, an extrapolation of program impacts to this population is equivalent to around 140,000-280,000 additional high school graduates.\(^{39}\) Moreover, the size of these program impacts is in the same range of magnitude as the effects found in similar CCT and education fee waiver programs in Colombia and Pakistan. The analysis also gives evidence of heterogeneity in program impacts in high school completion, with the effects being larger for girls and children in rural areas. In addition, the empirical analysis does not give conclusive evidence regarding the existence of indirect effects on the school completion rates of ineligible young adults in participating households.

5.3 By encouraging participant children to finish high school, *Familias en Acción* is expected to have other positive effects on further human capital gains, future employability, and income growth. High school completion may lead to significant payoffs in various dimensions. The first obvious channel is eligibility to apply for college or formal technical training, which may increase their qualifications and economic prospects. Indeed, recent evidence for Colombia shows that higher education (college and technical training) provides positive returns on wages that range from 7 to 18 percent. Likewise, having a high school diploma has value in the labor market in the form of improved access to more and better jobs and higher wages, by up to 7-11 percent (García-Suaza and others 2009, Prada 2006).\(^{40}\) Regarding other plausible dynamic effects, empirical findings in Colombia suggest that more educated individuals tend to not only have fewer children but also invest more in
their human capital. Conditional correlations show, for instance, that having a secondary education degree reduces the expected number of children by 27 percent, almost twice the effect calculated for primary education; in addition, children’s enrollment status and educational attainment are shown to be largely determined by their parents’ education (Forero and Gambo, 2009; Nunez and Sanchez 2003). It is important to note that while these estimates are based on other empirical research in the Colombian context, they do not pertain particularly to the program. Therefore, further work that tests these extrapolations, especially the program impacts on such outcomes as tertiary education, employment prospect, and long-term welfare, would provide important empirical evidence.

5.4 The expected benefits of increased school completion on future earnings reiterate the cost benefit analysis of the program provided by previous evaluations. A previous attempt to quantify the benefits of Familias en Acción simulates that, through improved infant health and increased years of secondary schooling, the program could increase future earnings and have benefits that are 1.59 times larger than the total costs (IFS and others 2006). By showing an actual increase in school attainment, the results of this paper shed additional light on the plausibility of the demonstrated cost-effectiveness of the program through the educational channel.

5.5 The impact analysis on learning outcomes shows mostly no statistical differences in test scores between the treated and equally disadvantaged control children, even after correcting for sample selection by excluding the “new enrollers” brought into school by the program. A first set of impact estimates, not corrected for selection bias, show that treated children perform as well as nonparticipant children in mathematics, Spanish, and the overall score of the ICFES exam. Since the beneficiaries used for this part of the analysis include both “new enrollers” and already enrolled children, this finding probably means that the former group improves their learning due to the program given that they perform as well as other poor children already in school. For the second set of estimates, restricting the treated sample to children that would have finished school even without the program to adjust for potential selection bias, there is still no clear pattern in the direction, size, and significance of effects on test scores across different model specifications. The distribution of impacts shows that the lack of effects is more evident for children in urban areas, while there is a set of results that point to a possible deterioration in the academic achievement of children that reside in rural municipalities.

5.6 The fact that children covered by Familias en Acción perform at the same level as equally poor nonparticipant children despite the money transfer and the conditionality is in line with existing evidence on similar programs in Mexico and Cambodia. As discussed in chapter 3, CCT programs like Familias en Acción could have various conflicting effects on the learning of participant children. While beneficiary children may be expected to perform better as they stay in school more and their parents invest more time and money in their nutrition, health and education, they may come from socio-economic background that, compared to nonbeneficiary children already in school, hinder their learning. Moreover, CCTs are often targeted to disadvantaged areas where the quality and supply of education are likely to be low. The increase in the demand for schooling could cause classroom congestion and induce negative peer effects. However, even though interventions like CCTs are not designed to directly raise learning, there is growing concern regarding the level of skills
with which program participants seek admission to higher education or enter the labor force after exiting the program. Therefore, assessing the potential of CCTs and/or supplementary interventions for increasing learning is critical. Policy makers could rethink some innovative design features such as changing the timing of transfers or tying them to performance rather than attendance. Supplementary supply-side interventions aimed at improving school quality and increasing resources for low-performing students are also possible options. Pilot tests, together with careful evaluations, would surely yield valuable knowledge about the efficacy of these policies to couple the objectives of increasing human capital with improving learning.
Annex A
Data Merging Procedures

Propensity Score Matching

The propensity score matching exercise builds on the short- and middle-term evaluation and uses survey data collected in 2002, 2003, and 2005 (three waves of panel data with more than 10,000 households each). Across the three waves, there are 51,056 individuals who were seven years or older at the baseline. This sample is merged with the ICFES database of around 6 million registrants between 2000 and 2009 to obtain their test scores and estimate whether they have completed high school (Figure 4). The ICFES tests are administered in Colombia twice a year, in October and May, and are often used as a criterion for enrollment in tertiary education. The majority of students who took the ICFES tests (90 percent) have finished grade 11 and around 90 percent of high school graduates take the test so the test registration is a good estimator for secondary school graduation (World Bank 1993; Angrist and others 2006). Although the ICFES exam takers in 2000 and 2001 could not have participated in Familias en Acción, which started in 2002, they are kept in the dataset because they may include older siblings of program participants for the analysis of program indirect effects.

Since one of the relevant outcomes is high school completion, different merging strategies are used to enhance the probability of finding the surveyed individuals in the ICFES database, ensuring that relevant people are not excluded. If an individual is not matched to any ICFES registration, it is assumed that the child did not take the test, either because they did not finish grade 11 or because they chose not to. The latter is more unlikely because of the high proportion of secondary school graduates that take the ICFES exam (Angrist and others 2006). There are many difficulties in matching the sampled children with the ICFES registration. However, there is no reason to believe that the matching errors, often due to name and ID mismatches, would be systematically different between the treatment and control groups.

Four different merging methods are employed:

- The first method uses only the ID numbers reported in the surveys as the matching criterion. This is a unique ID number assigned to all citizens of Colombia when they turn 11-12 years old. However, the ID numbers change when individuals turn 18, which may result in the failure to match individuals who took the ICFES test at or after 18 years of age. Furthermore, since an ID includes 11 digits, there are, expectedly, many occasions of IDs being misreported in the evaluation surveys. This merge, consequently, gives only 4,048 matched observations.

- The second method uses only full name as the matching criterion with the probability of orthographical mistakes (for instance, Catherine versus Katherine). While this may resolve the issue of ID change or misreporting as observed in the first method, it has some
potential mismatches due to a number of common last names in Colombia. It is therefore important to be cautious of the likelihood of matching different individuals with the same name. This merge results in 6,563 matches.

- The third method uses both the two last names and date of birth for merging. Again, due to many common last names in Colombia, this method does not guarantee unique matches. The merge provides 46,360 observations.

- The final method uses the two last names and the first 7 digits of the ID number for the merge. For children under 18 years of age, the first 6 digits correspond to the date of birth so this strategy is potentially more accurate than the third method of using last names and date of birth. With shorter IDs, it is also expected that there are fewer misreporting cases than in the first method. However, since the short ID numbers are not unique, this strategy has similar issue as the full name merge. It provides 5,927 matches.

The results from all four matches are used to minimize the probability of exclusion. In order to ensure that the matches are correct, the four merges are appended and subject to three cleaning processes. First, the records with the exact name and similar date of birth are kept (either the same date and month of birth and within four years of birth, or the same year of birth and within two months and two days of birth). Second, those that do not fulfill the first check are tested whether they have the exact date of birth and similar name (again, to account for orthographical mistakes). Finally, among those that fail both tests, the observations with similar name and similar date of birth are kept to account for mistakes in recording.

After each cleaning, duplicates are checked within both the evaluation survey sample and the ICFES dataset. Only the duplicates across both datasets are deleted to avoid eliminating individuals that took the exam multiple times. The result is 5,022 observations. When multiple test scores are found for one individual, the first test result is kept, which produces 4,820 records corresponding to unique individuals. This final data set makes up the sample used in the analysis.

The accuracy of the merges is tested using the information on students enrolled in schools in the evaluation surveys. This test follows the cohorts of students who in principle could have completed grade 11 between 2000 and 2009, assuming no grade repetition, and obtains 5,395 records. Again, this includes students who could have finished high school in 2000 and 2001 since they serve as the analysis group for another outcome. Incorporating the average dropout rates of students from grades 7 to 11, there should be 4,073 individuals who completed grade 11 within this time period. The final sample of 4,820 individuals obtained from the merging and cleaning process comes quite close to this estimate.

Among the 4,820 matched individuals, 3,002 have exactly the same full name and date of birth. Among the 5,022 matched records, 67 percent are 18 years old or younger and 88 percent are under 20 at the time they took the test. Most of the observations correspond to tests taken in the second semester of the year when public schools administer the exam, which is consistent with the fact that most of the program beneficiaries attend public schools.
Regression Discontinuity Design

The merge for the RD analysis uses three set of administrative data: (1) the System of Information of Beneficiaries of the program (SIFA) provided by Acción Social; (2) the Poverty Index Score Survey collected between 1994 and 2003 (SISBEN); and (3) records on registration for the national ICFES test (Figure 5). The following steps describe the use of these data:

- SIFA is used to construct the treatment groups while SISBEN provides the control group—individuals under and above the threshold of eligibility, respectively.

- After running the same merging procedures followed in the PSM analysis, the merge is able to identify that 95 percent of the matching distribution was born between the years of 1975 and 1994. This information is used to restrict the sample to those individuals that are mostly likely to be merged between SIFA+SISBEN and ICFES.

- The analysis focuses on information from the SISBEN survey that was collected between 1994 and 2003, since Familias en Acción targeted the beneficiaries during the first phase of the program with the scores from the first version of the poverty score index. For consistency of result, the sample is restricted in this way to evaluate comparable children who joined the program during the first expansion that took place between 2002 and 2003; these children come from the records of SIFA in 2001-06.

- The same four merging strategies employed in the PSM analysis were followed to guarantee the comparability of the outcome variable that measures secondary school completion.

- Individuals with score zero were excluded from the analysis for two reasons: First, it is not possible to establish whether a score zero is the result of a problem in the algorithm—the probability of getting 0 is very low. Second, the probability of getting Famí-
lia en Acción is much lower for this group when compared with the probability of receiving the treatment for people with scores equal to one.

Figure 5. Merging Method—RDD
### Descriptive Statistics

#### Table 7. Preprogram Summary Statistics by Treatment Status (Sample for the Matching Analysis)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treated</th>
<th>Control</th>
<th>Difference in means</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td><strong>n</strong></td>
<td><strong>Mean</strong></td>
<td><strong>n</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Demographic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (household head)</td>
<td>45.33</td>
<td>2,420</td>
<td>44.94</td>
<td>1,766</td>
</tr>
<tr>
<td>Age (spouse)</td>
<td>41.39</td>
<td>2,415</td>
<td>41.00</td>
<td>1,766</td>
</tr>
<tr>
<td>Age (child)</td>
<td>12.36</td>
<td>2,420</td>
<td>12.48</td>
<td>1,766</td>
</tr>
<tr>
<td>Gender (household head)</td>
<td>0.77</td>
<td>2,420</td>
<td>0.84</td>
<td>1,766</td>
</tr>
<tr>
<td>Gender (child)</td>
<td>0.44</td>
<td>2,420</td>
<td>0.43</td>
<td>1,766</td>
</tr>
<tr>
<td><strong>Household structure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is household single headed?</td>
<td>0.02</td>
<td>2,420</td>
<td>0.02</td>
<td>1,766</td>
</tr>
<tr>
<td>Number of household members</td>
<td>6.07</td>
<td>2,420</td>
<td>6.16</td>
<td>1,766</td>
</tr>
<tr>
<td>Number of children</td>
<td>1.34</td>
<td>2,420</td>
<td>1.42</td>
<td>1,766</td>
</tr>
<tr>
<td><strong>Consumption and assets</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly household consumption</td>
<td>199,747</td>
<td>2,362</td>
<td>212,330</td>
<td>1,697</td>
</tr>
<tr>
<td>Does the family own the house?</td>
<td>0.67</td>
<td>2,420</td>
<td>0.65</td>
<td>1,766</td>
</tr>
<tr>
<td><strong>Education, health and work</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does household head read?</td>
<td>1.00</td>
<td>2,415</td>
<td>1.00</td>
<td>1,762</td>
</tr>
<tr>
<td>Household head completed secondary or more?</td>
<td>0.06</td>
<td>2,267</td>
<td>0.08</td>
<td>1,673</td>
</tr>
<tr>
<td>Years of schooling (household head)</td>
<td>3.72</td>
<td>2,129</td>
<td>4.04</td>
<td>1,576</td>
</tr>
<tr>
<td>Did children suffer from diarrhea?</td>
<td>0.11</td>
<td>1,086</td>
<td>0.11</td>
<td>702</td>
</tr>
<tr>
<td>Does household head work?</td>
<td>0.89</td>
<td>2,316</td>
<td>0.89</td>
<td>1,720</td>
</tr>
<tr>
<td><strong>Dwelling characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Located in an urban area?</td>
<td>0.56</td>
<td>2,420</td>
<td>0.69</td>
<td>1,766</td>
</tr>
<tr>
<td>No walls?</td>
<td>0.01</td>
<td>2,414</td>
<td>0.01</td>
<td>1,765</td>
</tr>
<tr>
<td>Connected to piped water?</td>
<td>0.68</td>
<td>2,409</td>
<td>0.76</td>
<td>1,766</td>
</tr>
<tr>
<td>Connected to gas?</td>
<td>0.12</td>
<td>2,390</td>
<td>0.14</td>
<td>1,759</td>
</tr>
<tr>
<td>Connected to sewage?</td>
<td>0.33</td>
<td>2,417</td>
<td>0.32</td>
<td>1,766</td>
</tr>
<tr>
<td><strong>Community</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altitude</td>
<td>658.16</td>
<td>2,415</td>
<td>567.07</td>
<td>1,766</td>
</tr>
<tr>
<td>Students per teacher</td>
<td>22.49</td>
<td>2,415</td>
<td>22.68</td>
<td>1,766</td>
</tr>
<tr>
<td>Square meters of classroom per student</td>
<td>2.95</td>
<td>2,415</td>
<td>2.50</td>
<td>1,766</td>
</tr>
<tr>
<td>Number of banks</td>
<td>1.69</td>
<td>2,369</td>
<td>0.91</td>
<td>1,766</td>
</tr>
<tr>
<td>Number of health centers</td>
<td>1.13</td>
<td>2,369</td>
<td>0.84</td>
<td>1,766</td>
</tr>
<tr>
<td>Region = East?</td>
<td>0.21</td>
<td>2,415</td>
<td>0.24</td>
<td>1,766</td>
</tr>
<tr>
<td>Region = Central?</td>
<td>0.29</td>
<td>2,415</td>
<td>0.16</td>
<td>1,766</td>
</tr>
<tr>
<td>Region = Pacific?</td>
<td>0.11</td>
<td>2,415</td>
<td>0.13</td>
<td>1,766</td>
</tr>
<tr>
<td>Affected by violent attacks?</td>
<td>0.02</td>
<td>2,415</td>
<td>0.03</td>
<td>1,766</td>
</tr>
</tbody>
</table>

**Notes:** Significant at 90(*)*, 95(**), 99(***% confidence. Summary statistics calculated for households with at least one child (enrolled or not in school) who were 18 or below when the joined the program and that, based on their school attainment at the pre-program time, could have achieved grade 11 between 2003 and 2009.
Annex C
Models of Program Participation

Figure 6. Distribution of Propensity Scores on Prediction of Participation in *Familias en Acción* Used in Matching Estimation
### Table 8. Logit Estimates of the Determinants of Participation in *Familias en Acción* (Models I and II)

<table>
<thead>
<tr>
<th>Variables</th>
<th>(I)</th>
<th>(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.0130***</td>
<td>-0.0124***</td>
</tr>
<tr>
<td></td>
<td>[0.00430]</td>
<td>[0.00444]</td>
</tr>
<tr>
<td>Order of the child</td>
<td>-0.0384***</td>
<td>-0.0386***</td>
</tr>
<tr>
<td></td>
<td>[0.0145]</td>
<td>[0.0149]</td>
</tr>
<tr>
<td>Head of the household is married</td>
<td>0.162**</td>
<td>0.184**</td>
</tr>
<tr>
<td></td>
<td>[0.0718]</td>
<td>[0.0742]</td>
</tr>
<tr>
<td>Head of the household works</td>
<td>0.180***</td>
<td>0.156***</td>
</tr>
<tr>
<td></td>
<td>[0.0505]</td>
<td>[0.0521]</td>
</tr>
<tr>
<td>Male head of the household</td>
<td>-0.490***</td>
<td>-0.449***</td>
</tr>
<tr>
<td></td>
<td>[0.0735]</td>
<td>[0.0759]</td>
</tr>
<tr>
<td>Age head of the household</td>
<td>0.00171</td>
<td>0.00228</td>
</tr>
<tr>
<td></td>
<td>[0.00155]</td>
<td>[0.00160]</td>
</tr>
<tr>
<td>Urban</td>
<td>-0.495***</td>
<td>-0.398***</td>
</tr>
<tr>
<td></td>
<td>[0.0307]</td>
<td>[0.0335]</td>
</tr>
<tr>
<td>Years of schooling of the household head</td>
<td>0.00739</td>
<td>0.00459</td>
</tr>
<tr>
<td></td>
<td>[0.00593]</td>
<td>[0.00616]</td>
</tr>
<tr>
<td>Number of children ages 7 to 11</td>
<td>-0.0244</td>
<td>-0.0215</td>
</tr>
<tr>
<td></td>
<td>[0.0155]</td>
<td>[0.0159]</td>
</tr>
<tr>
<td>Number of children ages 12 to 17</td>
<td>-0.0153</td>
<td>-0.0263</td>
</tr>
<tr>
<td></td>
<td>[0.0156]</td>
<td>[0.0162]</td>
</tr>
<tr>
<td>Monthly expenditures</td>
<td>-8.50e-07****</td>
<td>-4.89e-07***</td>
</tr>
<tr>
<td></td>
<td>[1.56e-07]</td>
<td>[1.63e-07]</td>
</tr>
<tr>
<td>Teacher-pupil ratio in municipality</td>
<td>0.0222***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.00321]</td>
<td></td>
</tr>
<tr>
<td>Classroom space&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.0828***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.0101]</td>
<td></td>
</tr>
<tr>
<td>Resides in most dense part of municipality</td>
<td>-1.12e-05***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.02e-06]</td>
<td></td>
</tr>
<tr>
<td>Resides in least dense part of municipality</td>
<td>-7.74e-06***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[2.23e-06]</td>
<td></td>
</tr>
<tr>
<td>Number of urban schools registered in the municipality</td>
<td>0.0197***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.000906]</td>
<td></td>
</tr>
<tr>
<td>Number of rural Schools registered in the municipality</td>
<td>0.0331***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.00421]</td>
<td></td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>19,427</td>
<td>19,426</td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Coefficients do not correspond to marginal effects. Significant at 90(*) *, 95(**), 99(***) percent confidence. <sup>a</sup> Measured in square meters per student in the municipality.
### Table 9. Logit Estimates of the Determinants of Participation in *Familias en Acción* (Model III)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coeff.</th>
<th>s.e.</th>
<th>Variables</th>
<th>Coeff.</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live in a rural disperse area</td>
<td>0.390***</td>
<td>(0.0400)</td>
<td>Type of phone: shared or radiophone</td>
<td>0.206***</td>
<td>(0.0590)</td>
</tr>
<tr>
<td>Live in a rural populated area</td>
<td>0.782***</td>
<td>(0.0537)</td>
<td>No phone</td>
<td>0.643***</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Age household head</td>
<td>-0.00349*</td>
<td>(0.00212)</td>
<td>House is rented</td>
<td>0.0765</td>
<td>(0.0575)</td>
</tr>
<tr>
<td>Age spouse of household head</td>
<td>0.00227</td>
<td>(0.00229)</td>
<td>House occupied without property rights</td>
<td>-0.614***</td>
<td>(0.0677)</td>
</tr>
<tr>
<td>Marital status is single</td>
<td>0.0720</td>
<td>(0.0784)</td>
<td>Other type of property</td>
<td>-0.0863**</td>
<td>(0.0377)</td>
</tr>
<tr>
<td>Education of household head: incomplete primary</td>
<td>0.121***</td>
<td>(0.0368)</td>
<td>Number of violent deaths in municipality</td>
<td>0.234***</td>
<td>(0.0680)</td>
</tr>
<tr>
<td>Education of household head: complete primary</td>
<td>0.282***</td>
<td>(0.0523)</td>
<td>Number of violent attacks in municipality</td>
<td>-0.347***</td>
<td>(0.0773)</td>
</tr>
<tr>
<td>Education of household head: incomplete secondary</td>
<td>0.216***</td>
<td>(0.0657)</td>
<td>Two nuclear families living in the same house</td>
<td>0.536***</td>
<td>(0.0813)</td>
</tr>
<tr>
<td>Education of household head: complete secondary</td>
<td>0.0648</td>
<td>(0.102)</td>
<td>Three or more nuclear families living in the same house</td>
<td>0.956***</td>
<td>(0.161)</td>
</tr>
<tr>
<td>Education spouse: incomplete primary</td>
<td>0.0882**</td>
<td>(0.0381)</td>
<td>Oriental region</td>
<td>-1.255***</td>
<td>(0.0625)</td>
</tr>
<tr>
<td>Education spouse: complete primary</td>
<td>0.0338</td>
<td>(0.0514)</td>
<td>Central region</td>
<td>-0.152***</td>
<td>(0.0447)</td>
</tr>
<tr>
<td>Education spouse: incomplete secondary</td>
<td>-0.135**</td>
<td>(0.0658)</td>
<td>Pacific region</td>
<td>-0.906***</td>
<td>(0.0573)</td>
</tr>
<tr>
<td>Education spouse: complete secondary</td>
<td>-0.0542</td>
<td>(0.0982)</td>
<td>Altitud</td>
<td>0.000685***</td>
<td>(2.71e-05)</td>
</tr>
<tr>
<td>Family lives in a house or room</td>
<td>0.487***</td>
<td>(0.0999)</td>
<td>Resides in most dense part of municipality</td>
<td>-7.49e-06***</td>
<td>(1.92e-06)</td>
</tr>
<tr>
<td>Wall materials: Tapia, Abobe or Bahareque</td>
<td>0.00734</td>
<td>(0.0344)</td>
<td>Resides in least dense part of municipality</td>
<td>-1.22e-05***</td>
<td>(2.29e-06)</td>
</tr>
<tr>
<td>Wall materials: wood</td>
<td>-0.448***</td>
<td>(0.0478)</td>
<td>Number of urban schools registered in the municipality</td>
<td>0.0230***</td>
<td>(0.00400)</td>
</tr>
<tr>
<td>Wall materials: bad quality wood</td>
<td>0.257***</td>
<td>(0.0860)</td>
<td>Number of rural Schools registered in the municipality</td>
<td>0.0255***</td>
<td>(0.000870)</td>
</tr>
<tr>
<td>Wall materials: cardboard or no Walls</td>
<td>-0.168</td>
<td>(0.119)</td>
<td>Teacher pupil ratio in municipality</td>
<td>0.00841***</td>
<td>(0.00326)</td>
</tr>
<tr>
<td>House has is connected to natural gas</td>
<td>0.402***</td>
<td>(0.0602)</td>
<td>Classroom Square Meters per student in the municipality</td>
<td>0.103***</td>
<td>(0.00927)</td>
</tr>
<tr>
<td>House has water pipe</td>
<td>-0.0967***</td>
<td>(0.0341)</td>
<td>Observations</td>
<td>23,608</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Heteroskedasticity-robust standard errors clustered at the municipality level reported in parenthesis. Coefficients do not correspond to marginal effects. Significant at 90(*) , 95(**), 99(***)) percent confidence. * Measured in square meters per student in the municipality.
Annex D
Methodology for Nonparametric Bounds of Program Impacts

This paper uses the results of ICFES, a mandatory academic test given to students in school, to infer the impacts of Familias en Acción on learning outcome. This approach is problematic from a methodological standpoint because the program increases school enrollment and attendance among beneficiaries and with it the probability that they take the exam. Participant children are therefore more likely than nonparticipants to register for the ICFES test, which makes the scores distributions of beneficiaries and nonbeneficiaries not comparable. The intuition behind this is a selection bias created by the “new enroller” who is brought into school due to the incentives of the program and may be different in many dimensions (socioeconomic background, innate ability, motivation, expected returns to schooling, etc.) than those already enrolled in school. For these reasons, a simple comparison of participant and nonparticipant children may be deceptive and—given that the bias is expected to be negative—would probably underestimate the actual effect of the program on learning.

In order to address this identification issue, this evaluation uses bounding procedures on specific quantile average treatment effects estimated with matching techniques to correct for the selection bias brought about by the likely introduction of low-scorers into the group of program beneficiaries who ended up taking the test (Lee 2002; Angrist and others 2006). The two key assumptions of the procedure are: (1) independence of the treatment status and the errors in the outcome and selection equations (expected to be addressed by the PSM strategy); (2) a monotonicity condition in the sense that assignment to treatment affects the outcome in only one direction, namely that the program does not reduce the test scores of program participants:

\[ S_{1t} \geq S_{0t} \text{ for all } i \]

where \( S, 1 \) and \( i \) denote test scores, treatment status and individuals, respectively. Following Angrist and others (2006), students are assumed to choose to take the test if their expected scores are above a certain threshold so that the quantiles of test-takers are identified from the quantiles of non-takers (\( S_{tk} = 0 \) where \( t = 0,1 \) indexes treated and control children and \( k \) equals 1 if the student takes the test) for \( \varphi \geq \varphi_k \), where \( q_k(\varphi_k) = 0 \). The \( \varphi \)-quantile of the distributions of non-participants and participants are denoted by \( q_0(\varphi_0) \) and \( q_1(\varphi_1) \), respectively.

The main idea of the procedure is to find the quantile \( \varphi_0 \) for the control group such that \( q_0(\varphi_0) = 0 \) and restrict the distribution of \( S_{1} \) to the percentiles above \( \varphi_0 \). This defines the upper bound based on the subsample of individuals who would have taken the test regardless of the program \( E[S_{1} - S_{0} | k = 1] \) as follows:

\[ E[S | D = 1, S > q_0(\varphi)] - E[S | D = 0, S > q_0(\varphi)] \]
Unadjusted comparisons between treated and control children—conditional on positive test scores for each group—provide lower bounds on the actual effects of the program:

\[ E[S|D = 1, S > q_1(\varphi)] - E[S|D = 0, S > q_0(\varphi)] \]

Since the problem of selection is expected to be more prevalent at the bottom of the distribution, upper bounds should be tighter at upper parts of the distribution. Angrist (1997) shows that under the assumptions (1) and (2), the symmetric truncation of the score distributions is expected to eliminate the sample selection bias. As noted above, this equivalent to restricting the sample to the individuals thought to be the “always takers.”

The graphs below present the distribution of test scores in Spanish for the PSM sample without (panel a) and with correction for selection (panel b).

Notes: Test scores normalized by the mean and the standard deviation by year. Nonparametric bounds on the impacts estimated for students who would have taken the test even if they had not participated in the program. The units of observation are children (enrolled or not in school) who were 18 or below at baseline (2002) and that based on their school attainment at the preprogram time could have achieved grade 11 at any point between 2003 and 2009, the number of years needed to complete high school is lower than the number of years of treatment and registered for the ICFES test.
Annex E  
Distribution of Program Impacts

Table 10. OLS and Matching Estimates of Impacts on High School Completion  
(by gender and location)

<table>
<thead>
<tr>
<th>Outcome</th>
<th></th>
<th></th>
<th>OLS</th>
<th></th>
<th></th>
<th>Matching estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1. Sample: Boys</td>
<td></td>
<td></td>
<td>0.0080 0.0340</td>
<td>0.0206 0.0661*</td>
<td>0.0587*</td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td></td>
<td></td>
<td>[0.0260] [0.0260]</td>
<td>[0.0301] [0.0363]</td>
<td>[0.0348]</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,490</td>
<td>1,690</td>
<td>1,490</td>
<td>1,490</td>
<td>1,687</td>
<td></td>
</tr>
<tr>
<td>2. Sample: Girls</td>
<td></td>
<td></td>
<td>0.0440** 0.0580***</td>
<td>0.0523** 0.0856***</td>
<td>0.0899***</td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td></td>
<td></td>
<td>[0.0220] [0.0230]</td>
<td>[0.0245] [0.0290]</td>
<td>[0.0290]</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,986</td>
<td>2,198</td>
<td>1,986</td>
<td>1,986</td>
<td>2,198</td>
<td></td>
</tr>
<tr>
<td>3. Sample: Rural</td>
<td></td>
<td></td>
<td>0.0900*** 0.1030***</td>
<td>0.0868*** 0.1044***</td>
<td>0.1176***</td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td></td>
<td></td>
<td>[0.0280] [0.0290]</td>
<td>[0.0314] [0.0402]</td>
<td>[0.0398]</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,356</td>
<td>1,516</td>
<td>1,356</td>
<td>1,356</td>
<td>1,514</td>
<td></td>
</tr>
<tr>
<td>4. Sample: Urban</td>
<td></td>
<td></td>
<td>-0.0100 0.0140</td>
<td>-0.0052 0.0492*</td>
<td>0.0391</td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td></td>
<td></td>
<td>[0.0210] [0.0220]</td>
<td>[0.0229] [0.0274]</td>
<td>[0.0254]</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,120</td>
<td>2,372</td>
<td>2,120</td>
<td>2,120</td>
<td>2,372</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Bootstrapped standard errors reported in square brackets are obtained from 200 replications. Significant at 90(*), 95(**), 99(***). Units of observation are children (enrolled or not in school) who were 18 or below at baseline (2002) and that, based on their school attainment at the pre-program time, could have achieved grade 11 between 2003 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Units of analysis are matched on the propensity score from three different specifications of a logistic regression on participation in the program. Preprogram covariates of each specification of the logit models of participation are listed in Table 8 and Table 9 in Annex C. OLS coefficients in models 1 and 2 are estimated from specifications without any covariates and with all the control variables included in Model 3 of the matching analysis, respectively.
ANNEX E
DISTRIBUTION OF PROGRAM IMPACTS

Figure 7. RDD Estimates of Impacts on High School Completion (by gender and area)

Table 11. RDD Reduced Form Estimates of Impacts on High School Completion (by gender and area)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Quadratic</th>
<th>Cubic</th>
<th>Quartic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1. Sample: Boys</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>0.010**</td>
<td>0.018***</td>
<td>0.023***</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.004]</td>
<td>[0.005]</td>
</tr>
<tr>
<td>Observations</td>
<td>401,989</td>
<td>401,989</td>
<td>401,989</td>
</tr>
<tr>
<td>2. Sample: Girls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>0.001</td>
<td>0.010**</td>
<td>0.029***</td>
</tr>
<tr>
<td></td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>Observations</td>
<td>414,049</td>
<td>414,049</td>
<td>414,049</td>
</tr>
<tr>
<td>3. Sample: Urban</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.023***</td>
</tr>
<tr>
<td></td>
<td>[0.005]</td>
<td>[0.005]</td>
<td>[0.006]</td>
</tr>
<tr>
<td>4. Sample: Rural</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted FA</td>
<td>0.026***</td>
<td>0.027***</td>
<td>0.022***</td>
</tr>
<tr>
<td></td>
<td>[0.004]</td>
<td>[0.004]</td>
<td>[0.005]</td>
</tr>
<tr>
<td>Observations</td>
<td>408,979</td>
<td>408,979</td>
<td>408,979</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***).
Table 12. Upper Bounds on Program Effects on Test Scores by Gender and Area (matching estimates)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mathematics</th>
<th>Spanish</th>
<th>Overall test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>All</td>
<td>Effect</td>
<td>0.033</td>
<td>0.0374</td>
</tr>
<tr>
<td></td>
<td>s.e.</td>
<td>[0.0419]</td>
<td>[0.0495]</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>1,850</td>
<td>1,850</td>
</tr>
<tr>
<td>Boys</td>
<td>Effect</td>
<td>0.0443</td>
<td>-0.0028</td>
</tr>
<tr>
<td></td>
<td>s.e.</td>
<td>[0.0762]</td>
<td>[0.0868]</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>722</td>
<td>722</td>
</tr>
<tr>
<td>Girls</td>
<td>Effect</td>
<td>0.0451</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>s.e.</td>
<td>[0.0596]</td>
<td>[0.0691]</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>1,119</td>
<td>1,119</td>
</tr>
<tr>
<td>Rural</td>
<td>Effect</td>
<td>0.2735***</td>
<td>0.2856***</td>
</tr>
<tr>
<td></td>
<td>s.e.</td>
<td>[0.0851]</td>
<td>[0.0998]</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>557</td>
<td>557</td>
</tr>
<tr>
<td>Urban</td>
<td>Effect</td>
<td>0.0246</td>
<td>0.0514</td>
</tr>
<tr>
<td></td>
<td>s.e.</td>
<td>[0.0565]</td>
<td>[0.073]</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>1,219</td>
<td>1,219</td>
</tr>
</tbody>
</table>

Notes: Test scores are normalized by the mean and the standard deviation in each subject by year. The definition of the overall test score excludes results of the foreign language test chosen by the student. Nonparametric bounds on the impacts estimated for students who would have taken the test even if they had not participated in the program. Percentile in first row of the table corresponds to the percentile φ in which where selection-free control children begin having positive scores. Bootstrapped standard errors reported in square brackets are obtained from 200 replications. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children enrolled or not in school who were 18 or below at baseline (2002) and that based on their school attainment at the preprogram time could have achieved grade 11 at any point between 2003 and 2009, the number of years needed to complete high school was lower than the number of years of treatment and registered for the ICFES test. Units of analysis matched on the propensity score from three different specifications of a logistic regression on participation in the program. Preprogram covariates of each specification of the logit models of participation are listed in Table 8 and Table 9 in Annex C.

Table 13. RDD Reduced Form Estimates of Impacts on Mathematics, Spanish, and Overall Test Scores (by gender and area)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mathematics</th>
<th>Spanish</th>
<th>Overall test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys</td>
<td>Girls</td>
<td>Urban</td>
</tr>
<tr>
<td>Predicted FEA</td>
<td>0.019</td>
<td>0.041</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>[0.039]</td>
<td>[0.033]</td>
<td>[0.037]</td>
</tr>
<tr>
<td>Observations</td>
<td>92,054</td>
<td>114,258</td>
<td>129,389</td>
</tr>
<tr>
<td>Functional Form</td>
<td>Quartic</td>
<td>Quartic</td>
<td>Quartic</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>Yes</td>
<td>yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: Scores are normalized by the mean and the standard deviation in each subject by year. The definition of the overall test score excludes results of the foreign language test chosen by the student. Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children enrolled or not in school who were 18 or below at baseline (2002) and that based on their school attainment at the preprogram time could have achieved grade 11 at any point between 2003 and 2009, the number of years needed to complete high school was lower than the number of years of treatment and registered for the ICFES test. Models include quartic [(Si - S*)4] specifications of the control function below and above the cutoff of eligibility S*. 

51
Annex F
Estimates of Program Indirect Effects

Table 14. OLS and Matching Estimates of Impacts on the High School Completion of Nonparticipating Young Adults (total and by gender and area)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>Matching estimates</td>
<td>OLS</td>
<td>Matching estimates</td>
<td>OLS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Sample: All</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td>0.043***</td>
<td>0.044***</td>
<td>0.0386***</td>
<td>0.0434***</td>
<td>0.0268</td>
</tr>
<tr>
<td></td>
<td>[0.014]</td>
<td>[0.014]</td>
<td>[0.0142]</td>
<td>[0.0176]</td>
<td>[0.0185]</td>
</tr>
<tr>
<td>Observations</td>
<td>2,300</td>
<td>2,300</td>
<td>2,300</td>
<td>2,300</td>
<td>2,321</td>
</tr>
<tr>
<td>2. Sample: Boys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td>0.039**</td>
<td>0.040**</td>
<td>0.0424**</td>
<td>0.0491***</td>
<td>0.0272</td>
</tr>
<tr>
<td></td>
<td>[0.018]</td>
<td>[0.018]</td>
<td>[0.0167]</td>
<td>[0.0186]</td>
<td>[0.0280]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,417</td>
<td>1,417</td>
<td>1,417</td>
<td>1,417</td>
<td>1,481</td>
</tr>
<tr>
<td>3. Sample: Girls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td>0.044*</td>
<td>0.049**</td>
<td>0.0389</td>
<td>0.0415</td>
<td>0.0245</td>
</tr>
<tr>
<td></td>
<td>[0.024]</td>
<td>[0.025]</td>
<td>[0.0248]</td>
<td>[0.0313]</td>
<td>[0.0369]</td>
</tr>
<tr>
<td>Observations</td>
<td>883</td>
<td>883</td>
<td>883</td>
<td>883</td>
<td>840</td>
</tr>
<tr>
<td>4. Sample: Rural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td>0.040**</td>
<td>0.042**</td>
<td>0.0321</td>
<td>0.0214</td>
<td>0.0399</td>
</tr>
<tr>
<td></td>
<td>[0.019]</td>
<td>[0.020]</td>
<td>[0.0200]</td>
<td>[0.0310]</td>
<td>[0.0272]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,117</td>
<td>1,117</td>
<td>1,117</td>
<td>1,117</td>
<td>1,135</td>
</tr>
<tr>
<td>5. Sample: Urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school completion</td>
<td>0.044**</td>
<td>0.039*</td>
<td>0.0381</td>
<td>0.0414</td>
<td>0.0163</td>
</tr>
<tr>
<td></td>
<td>[0.021]</td>
<td>[0.021]</td>
<td>[0.0223]</td>
<td>[0.0258]</td>
<td>[0.0329]</td>
</tr>
<tr>
<td>Observations</td>
<td>1,183</td>
<td>1,183</td>
<td>1,183</td>
<td>1,183</td>
<td>1,179</td>
</tr>
</tbody>
</table>

Notes: Bootstrapped standard errors reported in square brackets are obtained from 200 replications. Significant at 90(*), 95(**), 99(***). The units of observation are individuals (enrolled or not in school) who were 18 years or above at baseline (2002), live in a household with at least one child eligible for the program and that, based on their school attainment at the pre-program time, could have achieved grade 11 between 2003 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Units of analysis are matched on the propensity score from three different specifications of a logistic regression on participation in the program. Preprogram covariates of each specification of the logit models of participation are listed in Table 8 and Table 9 in Annex C. OLS coefficients in (1) and (2) are estimated from specifications without any covariates and with all the control variables included in model III of the matching analysis, respectively. CS stands for common support.
### Table 15. RDD Reduced Form Estimates of Impacts on the High School Completion of Nonparticipating Young Adults

<table>
<thead>
<tr>
<th>Variable</th>
<th>Quadratic</th>
<th>Cubic</th>
<th>Quartic</th>
<th>Non-parametric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Predicted FA</td>
<td>0.023</td>
<td>0.032</td>
<td>0.052</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>[0.026]</td>
<td>[0.027]</td>
<td>[0.037]</td>
<td>[0.035]</td>
</tr>
<tr>
<td>Year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Municipality fixed effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>14681</td>
<td>14668</td>
<td>14681</td>
<td>14668</td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(∗), 95(∗∗), 99(∗∗∗) percent confidence. The units of observation are individuals (enrolled or not in school) who were 18 years or above at baseline (2002), live in a household with at least one children eligible for the program and that, based on their school attainment at the pre-program time, could have achieved grade 11 between 2002 and 2009 and the number of years needed to complete high school was lower than the number of years of treatment. Models include linear [(Si - S*)], quadratic [(Si - S*)2], cubic [(Si - S*)3], and quartic [(Si - S*)4] specifications of the control function below and above the cutoff of eligibility S*. 
Annex G
Robustness Checks

Figure 8. Continuity Checks for Household- and Individual-Level Variables
### Table 16. Continuity Checks for Preprogram Household- and Individual-Level Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Within Band of 1 point</th>
<th>0.25 points</th>
<th>Variable</th>
<th>Within Band of 1 point</th>
<th>0.25 points</th>
<th>Variable</th>
<th>Within Band of 1 point</th>
<th>0.25 points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Household Variables</strong></td>
<td></td>
<td></td>
<td><strong>Panel B: Individual Variables</strong></td>
<td></td>
<td></td>
<td><strong>Panel C: Characteristics of the house</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income per capita</td>
<td>-508.67</td>
<td>5,655</td>
<td>Male</td>
<td>0.052</td>
<td>0.052</td>
<td>Access to electricity</td>
<td>-0.022</td>
<td>-0.138</td>
</tr>
<tr>
<td></td>
<td>[1,065.3]</td>
<td>[2,620.8]</td>
<td></td>
<td>[0.036]</td>
<td>[0.082]</td>
<td></td>
<td>[0.048]</td>
<td>[0.114]</td>
</tr>
<tr>
<td>Total income</td>
<td>84.107</td>
<td>8,739.60</td>
<td>Age</td>
<td>-0.398*</td>
<td>-0.103</td>
<td>Access to sewage</td>
<td>0.068</td>
<td>0.458**</td>
</tr>
<tr>
<td></td>
<td>[5,802.3]</td>
<td>[14,549.6]</td>
<td></td>
<td>[0.219]</td>
<td>[0.487]</td>
<td></td>
<td>[0.104]</td>
<td>[0.230]</td>
</tr>
<tr>
<td>Head of household earnings</td>
<td>-2,308.13</td>
<td>1,630.80</td>
<td>Levels of schooling</td>
<td>-0.062</td>
<td>0.018</td>
<td>Access to water</td>
<td>-0.082</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>[4,882.5]</td>
<td>[12,849.3]</td>
<td></td>
<td>[0.040]</td>
<td>[0.089]</td>
<td></td>
<td>[0.071]</td>
<td>[0.158]</td>
</tr>
<tr>
<td>Household size</td>
<td>0.009</td>
<td>0.567*</td>
<td>Regular activity</td>
<td>0.071</td>
<td>0.179</td>
<td>Time to get water</td>
<td>-0.065</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>[0.155]</td>
<td>[0.330]</td>
<td></td>
<td>[0.160]</td>
<td>[0.361]</td>
<td></td>
<td>[0.046]</td>
<td>[0.102]</td>
</tr>
<tr>
<td>Number of children under six</td>
<td>0.101</td>
<td>0.358**</td>
<td>School enrollment</td>
<td>-0.017</td>
<td>-0.029</td>
<td>Trash disposal service</td>
<td>-0.126**</td>
<td>-0.188</td>
</tr>
<tr>
<td></td>
<td>[0.098]</td>
<td>[0.180]</td>
<td></td>
<td>[0.032]</td>
<td>[0.072]</td>
<td></td>
<td>[0.057]</td>
<td>[0.123]</td>
</tr>
<tr>
<td>Schooling -- household head</td>
<td>0.262***</td>
<td>0.202</td>
<td>Type of school</td>
<td>-0.070*</td>
<td>0.081</td>
<td>Home ownership</td>
<td>0.122**</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td>[0.082]</td>
<td>[0.178]</td>
<td></td>
<td>[0.042]</td>
<td>[0.097]</td>
<td></td>
<td>[0.059]</td>
<td>[0.132]</td>
</tr>
<tr>
<td>Male head of the household</td>
<td>-0.031</td>
<td>-0.056</td>
<td>Farmer</td>
<td>0.005</td>
<td>0.013</td>
<td>Own a refrigerator</td>
<td>0.029</td>
<td>0.129**</td>
</tr>
<tr>
<td></td>
<td>[0.029]</td>
<td>[0.064]</td>
<td></td>
<td>[0.005]</td>
<td>[0.011]</td>
<td></td>
<td>[0.023]</td>
<td>[0.050]</td>
</tr>
<tr>
<td>Age of household head</td>
<td>-1.954**</td>
<td>0.567</td>
<td>Marital status</td>
<td>0.001</td>
<td>0.023</td>
<td>Own a television</td>
<td>-0.036</td>
<td>0.332***</td>
</tr>
<tr>
<td></td>
<td>[0.842]</td>
<td>[1.901]</td>
<td></td>
<td>[0.008]</td>
<td>[0.015]</td>
<td></td>
<td>[0.032]</td>
<td>[0.073]</td>
</tr>
<tr>
<td>Schooling -- spouse</td>
<td>0.172*</td>
<td>0.191</td>
<td></td>
<td></td>
<td></td>
<td>Own a Fan</td>
<td>-0.023</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>[0.093]</td>
<td>[0.208]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.019]</td>
<td>[0.043]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Number of bedrooms</td>
<td>-0.082</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.059]</td>
<td>[0.130]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Living room in the house</td>
<td>-0.040</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.027]</td>
<td>[0.061]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dining room in the house</td>
<td>-0.024</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.018]</td>
<td>[0.043]</td>
</tr>
<tr>
<td>Observations</td>
<td>47,205</td>
<td>12,130</td>
<td></td>
<td></td>
<td></td>
<td>Observations</td>
<td>47,205</td>
<td>12,130</td>
</tr>
<tr>
<td>Number of Municipalities</td>
<td>476</td>
<td>465</td>
<td></td>
<td></td>
<td></td>
<td>Number of Municipalities</td>
<td>476</td>
<td>465</td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors clustered at the municipality level reported in square bracket. Significant at 90(*), 95(**), 99(***) percent confidence. Regressions include quartic specifications of the control function below and above the cutoff of eligibility $S^*$. 

---

### ANNEX G

**ROBUSTNESS CHECKS**

---

56
Figure 9. Distribution of the SISBEN Score (total and by gender and area)
References


Endnotes

1. Earlier evaluations show that FA has modest increases in school enrollment, particularly for younger beneficiaries. The improvement is only 2 percentage points for children in primary school in rural areas and 5-7 percentage points for children in secondary school.

2. Subsidios is another CCT in Colombia, implemented mainly in Bogota. While the transfer amount and conditions of the program are similar to FA, Subsidios includes two other treatments with different benefit structures. One is the “savings treatment,” which reserves one third of the bimonthly payment to be given to families as a lump sum at the end of the year, just before enrollment into the subsequent grade level. The other is the “tertiary treatment,” which also reserves part of the bimonthly payment, but then pays families a substantially larger amount if students graduate from high school (grade 11). Students who continue to tertiary education are eligible to receive that amount one year earlier than those who do not. These two additional treatments may change the way families respond to the program in two ways: one is due to the reduction in the amount of cash families receive on a bimonthly basis, therefore limiting their liquidity; another is through the incentives created by linking cash transfers directly to grade progression and matriculation in tertiary education.

3. It is worth noting that the results of these two papers are not generalizable to the entire population covered by Oportunidades given that the samples of analysis only contain non-migrant rural households. Migrants are expected to have higher enrollment in college and better job opportunities, including higher salaries.

4. Results from Behrman and Parker (2008) actually show some positive trends in math and reading tests for program participants. Results on math achievement tests are however limited by low sample sizes while the analysis on reading tests is based on single differences between program participants which according to the authors could underestimate the actual impacts of the program.

5. The only known exceptions include a study that tracked participants of Oportunidades in Mexico 5-6 years into the program (Behrman and others 2005), an evaluation of a private secondary school voucher program in Colombia 5-6 years after the program started (Angrist and others 2004), and a paper on a female secondary school stipend program in Pakistan 4 years into implementation (Alam and others 2010).

6. Secondary level includes grades 6-9 and high school level includes grades 10-11.

7. The unique individual identification numbers were used to reduce duplications.

8. The list of schools serving participant children is confirmed annually.

9. Except for variation in the time allowed to register participants, the two empirical analysis use proxy variables to control for these factors either directly (as in the PSM approach) or indirectly (as captured in the score of the proxy-means test for the RDD).

10. Municipalities were grouped based on the number of eligible families that reside in each of them to form 639 Primary Sampling Units (PSU). Twenty-five strata were then defined based on geographic location, level of urbanization, number of eligible families, and indexes of quality of life and availability of school and health facilities in the municipality. Fifty PSUs (two within each stratum) corresponding to 57 municipalities targeted by the program were selected and then matched to 50 “control” PSUs (equivalent to 9 municipalities) that were relatively similar to “treatment” PSUs. Matching was done within each stratum and the comparability of both types of PSUs was assessed on the basis of population size and an index of quality of life. Finally, a stratified random sample of eligible households was selected in treatment and matched control municipalities. For more details about the matching process and the stratified random methods followed in the evaluation, see Attanasio and others (2005).

11. It is important to caution that the baseline measurement is not strictly a pre-program picture of all treated households because the program had just begun in a number of treatment municipalities at the time the baseline survey was being collected. To account for this, the treatment sample was divided into two groups each with 25 PSUs. One group consisted of municipalities where the program had already started (labeled as TCP which
stands for “tratamiento con pago”) and the other group of municipalities where the program had not been implemented (labeled as TSP, which stands for “tratamiento sin pago”) but program awareness campaigns and registration had started.

12. Despite the fact that the presence of cohorts would induce variation of treatment dosage, the evaluation design used in this paper cannot disentangle the effects of age on the outcomes from those of length of exposure. Conditional on being enrolled in school, older beneficiaries (when they joined the program) have fewer years ahead of them in school, and perhaps, are more likely to be observed finishing high school. If this is the case, shorter length of exposure may be wrongly attributed to higher school completion rates.

13. Common support rates for most impact estimates obtained with matching methods are above 97 percent. The lowest common support rate (92 percent) was for models of school completion in the sample with only rural households.

14. Results from this test are also required to go to college.

15. Only a very small fraction of children were found to have taken the test more than once. However, although the test can be taken several times, it is administered only to students who have achieved grade 11.

16. See notes to the tables with econometric results for more details.

17. This phase covered 622 municipalities and started with a few pilot municipalities in 2001.

18. Even though the information system of the program also reports SISBEN scores for participant families, the scores calculated by the SISBEN algorithm are used for them in order to apply a uniform score for the whole sample.


20. Nearly 12 percent of the households interviewed in treatment municipalities were not registered with the program (Attanasio and others 2005).

21. Around 95 percent of the children matched to ICFES records are in this age range.

22. In order to study the indirect effects of participant children on ineligible young adults that were still in school or could rejoin, this sample also includes individuals 18 to 23 years old that were listed as dependents.

23. A new SISBEN survey was fielded between 2003 and 2005 to update the information, improve the effectiveness of the targeting scheme, and change the algorithm due to concerns regarding manipulation by local authorities (Barrera, Linden, and Urquiola 2007).

24. The final sample also excludes internally displaced people who became eligible to the program much later.

25. The Levenshtein distance measures the difference between two strings as the number of edits (that is, the number of single characters) needed to transform one string into the other. For example, the Levenshtein distance between the names Yonatan and Jonathan is two as that is the exact number of characters that have to be changed (added, subtracted, inverted with adjacent characters, or modified) to convert one name into the other: from Yonatan to Jonatan and then to Jonathan.

26. A small fraction of individuals also take the ICFES test after they have finished high school. Also, since the test can be taken many times, only the first registration date and score is kept for the small number of students who took it more than once.

27. Due to small variations in the test questions, formats and scoring systems across years, ICFES advised normalizing the test scores by year and subject. Another normalization (often used by ICFES) that penalizes scores far away from the mean was also used and the results are almost identical. Although normalization based on a pre-program year might help eliminate some biases that may be induced by the composition and crowding effects resulting from newly enrolled students, it could introduce other biases due to the differences among tests across years.

28. Information to disentangle these effects would be useful for improving the design of the program. Unfortunately, the available data do not allow such separation.

29. Evidence of such mechanism has been found for Mexico’s Oportunidades, where program participants in secondary school became less likely to complete grades. This empirical fact is a possible consequence of the disin-
centives provided by the program termination after the third year of secondary school and because students want to remain as long as possible in the program (Dubois and others 2004).

30. Results presented in the text were obtained from Epanechnikov kernels. Other types of kernels used in the analysis include gaussian, tricube, and matching on the five closest neighbors.

31. A similar CCT program implemented in Bogota (Colombia), Subsidios, increased the probability of completing high school by 4 percentage points. Beneficiaries of Colombian PACES, which offered vouchers to attend private secondary schools to students from poor urban neighborhoods, are also found to be 5-7 percentage points more likely to graduate from high school. In Pakistan, the Female School Stipend program, a CCT targeting girls, appears to improve the chance of completing grade 9 in high school for girls aged 15-16 years old by 5 percentage points (Angrist and others 2006; Barrera-Osorio and others 2008; Alam and others 2010).

32. The cutoff of eligibility takes different values for urban and rural municipalities. Therefore, rather than presenting the outcomes as a function of the SISBEN, they are normalized as the distance of each child’s score to the area-specific cutoff that is used to classify households as SISBEN 1 and determine eligibility to the program. For instance, a child with a value -5 is in reality 5 points below the cutoff and is therefore eligible to Familias en Acción. In contrast, positive values of the normalized score represent children that belong to ineligible households.

33. The base specification includes quadratic forms of the control function (Si-S*), where Si correspond to the value of the proxy-means test and S* is the threshold of eligibility. Furthermore, regression functions are allowed to differ on both sides of the cutoff point (models I and II). Other flexible specifications of the regression function are also estimated including cubic and quartic functional forms of the control function (models III to VI).

34. In fact, cross-country evidence shows that school attainment correlates strongly with test scores (Filmer and others 2006).

35. This analysis, focusing only on children finishing high school, might underestimate program impacts on learning due to attrition. That is, it does not account for the improved learning of participant children who are brought into school by the program but do not finish high school and/or take the ICFES test.

36. Program effects are also higher for girls when the sample is broken down by location (urban versus rural).

37. The final sample includes young adults that meet the criteria outlined in the text and that were identified in the survey as household members but not as the main breadwinner.

38. It was not possible to control for school fixed effects in the econometric analysis for at least three reasons. First, high school completion is observed only for individuals in the relevant ages that registered for the ICFES test and therefore it is impossible to know the school in which children that are not matched to the ICFES database were enrolled. Second, information with the name and code of schools exists only for a—probably not random—subsample of all the students who took the ICFES test. Third, in many cases—particularly in small municipalities where the school supply is fixed—the inclusion of municipality fixed effects removes any existing school fixed effects. Furthermore, and related to this, the sample used in the RD approach includes a groups of children above and below the cutoff from the same municipality who are likely to go to the same school.

39. If households classified in SISBEN 1 have on average 1.37 children between 7 and 18 years and the number of families covered by the program is about 2.8 million, FA supports nearly 3.8 million children. Based on the high school completion rates of the control group from the PSM analysis, approximately 50 percent of these children would graduate from high school. The effect of the program in terms of the number of students is proxied by the increment in the high school completion rates induced by the program (from 50 percent to 54-58 percent) with respect to the 3.8 million that are currently participating in FA.

40. These future education and employment payoffs, of course, depend largely on factors on the supply side, that is, the ability of the Colombian tertiary education system and labor market to absorb this influx.

41. The IDs distributed before 2004 have 9 digits. We used 7 digits in order to maximize the number of merges we could get.