LAURA ABRAMOVSKY
Institute for Fiscal Studies

ORAZIO ATTANASIO
University College London
and Institute for Fiscal Studies

KAI BARRON
University College London

PEDRO CARNEIRO
University College London,
Institute for Fiscal Studies, and Centre
for Microdata Methods and Practice

GEORGE STOYE
Institute for Fiscal Studies
and University College London

Challenges to Promoting Social Inclusion of the Extreme Poor: Evidence from a Large-Scale Experiment in Colombia

ABSTRACT We evaluate the large-scale pilot program of an innovative and major welfare intervention in Colombia, which combines home visits by trained social workers to households in extreme poverty with preferential access to social programs. We use a randomized control trial and a very rich data set collected as part of the evaluation to identify program impacts on the knowledge and take-up of social programs and the labor supply of targeted households. We find no consistent impact of the program on these outcomes, possibly because the way the pilot was implemented resulted in very light treatment in terms of home visits. Importantly, administrative data indicate that the program has been rolled out nationally in a very similar fashion, suggesting that this major national program is likely to fail in making a significant contribution to reducing extreme poverty. We suggest that the program should undergo substantial reforms, which in turn should be evaluated.

JEL classification: J08, I38, C93

Keywords: Social exclusion, Social protection, Colombia, Extreme poverty, Labor supply, Randomized control trial

---

ACKNOWLEDGMENTS *Economía* thanks LACEA's Labor Network for financial support to cover the publication costs of this article.

The authors would like to thank Samuel Freije for helpful suggestions and Marcos Vera-Hernández, Emla Fitzsimons, and Susana Martínez-Restrepo for helpful comments. They are grateful to Juan Camilo Mejía, Victor Hugo Zuluaga, and researchers at Sistemas Especializados de Información (SEI) and Econometría Consultores in Colombia for help with collecting and assisting with the data. They gratefully acknowledge financial support provided by the Economic and Social Research Council under the Center for the Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (grant number RES-544-28-0001), the International Development Research Center from Canada, and the European Research Council grant 249612 on "Exiting Long-Run Poverty: The Determinants of Asset Accumulation in Developing Countries." They are also grateful to the Government of Colombia, which provided funding for the initial data collection and evaluation analysis via its National Planning Department.

Households that live in extreme poverty often face a multitude of interacting constraints that prevent them from improving their lives.[1] The causes of poverty traps have been much debated, and yet it is not always obvious where market imperfections and frictions arise, nor how to effectively tackle them. Existing work emphasizes capital and skill constraints as an important mechanism leading to the persistence of poverty.[2] Related literature further highlights coordination problems and psychological and behavioral constraints that arise due to poverty.[3] To address these issues, countries commonly set up a range of social programs, usually aimed at addressing one constraint at a time. However, many of the individuals that are most likely to benefit are often the least likely to enroll in such programs, perhaps because of a lack of knowledge, stigma, overly complex programs, and a lack of self-control.[4] This paper evaluates a large-scale social program in Colombia that aims to address these issues.

In 2007, the Colombian government launched a large-scale pilot program called *Juntos,* designed to tackle extreme poverty. This program aimed to address a number of different monetary and nonmonetary constraints to improve economic outcomes and the welfare of the poorest families along a number of dimensions, including improvements in health, housing, nutrition, and labor outcomes. The main objectives of the program were to build in indigent families the basic capacities to sustainably manage their own development and to stimulate demand for existing social programs. The program attempted to achieve these goals through home visits from social workers over a five-year period, as well as the expansion and improvement of the supply of existing programs in a coordinated effort by federal, regional, and local government agencies. It was rolled out on a national scale in the second half of 2011 under the name *Unidos,* with broadly the same scheme and aims.[5] The national program now targets 1.5 million families and accounted for 5 percent of the total public budget for social inclusion in 2013.[6] This program was inspired by *Chile Solidario,* which was introduced in Chile in 2002.

1. See, for example, Duflo (2012).

2. For example, Banerjee and Newman (1993); Galor and Zeira (1993); Ghatak and Jiang (2002).

3. On coordination problems, see Kremer (1993); on psychological and behavioral constraints, see Mullainathan and Shafir (2013); Dalton, Ghosal, and Mani (2014).

4. See Currie (2006).

5. Since the data that we analyze pertain to the initial phase of the program, we refer to the program as *Juntos* except when specifically referring to the current program in Colombia.

6. See Ministry of Finance and Public Credit, "Mensaje Presidencial: Proyecto de Presupuesto General De la Nacion, 2013," table 31. *Familias en Acción* accounts for almost 40 percent of this budget, serving around 2.2 million poor families.

Programs similar in nature to *Unidos* have become increasingly popular as a core strategy to alleviating poverty in a number of other Latin American countries, including Brazil, Mexico, and Peru.[7] Understanding the impacts of this program is therefore a high priority for policymakers across a number of countries.

In this paper, we examine the short-run impact of *Juntos* on the knowledge of a range of existing social programs, the take-up of the main Colombian conditional cash transfer program (*Familias en Acción*), and labor market outcomes.[8] These labor market outcomes include the participation rate, employment rate (and type of employment), unemployment rate, and hours worked, as well as employment earnings and tenure. Impacts cover the initial eighteen-month period following the implementation of the program across three main groups within the extreme poor in Colombia: rural, urban, and displaced households. The impact results are estimated using a large data set collected as part of a large randomized control trial. We observe some selection into the treatment group and into the panel sample that could potentially be nonrandom, especially for the urban population. We carefully document this for the three different representative samples and provide difference-in-differences intention-to-treat (ITT) and instrumental variable (IV) estimates that aim to correct for these potential biases.

Our results suggest that *Juntos* had no systematic or significant effects on the outcomes of interest. For example, in rural households, we find no impact on the knowledge of existing social programs. We find a positive impact on the use of *Familias en Acción,* which is relatively large in magnitude but only statistically significant at the 10 percent level. We find no positive impact on labor market outcomes. In fact, we find negative effects on the probability of employment for rural women, driven by a decrease in self-employment within this group. This is accompanied by a decrease in the hourly pay for rural women. These results are consistent with recent empirical evidence on conditional cash transfers, which suggests that these programs are associated with a decrease in the labor supply of beneficiary individuals, especially among women with young children.[9] However, given the large number of

7. Chile was the first country to introduce this type of program, with the implementation of *Chile Solidario* in 2002. Brazil introduced a similar program called Brasil sem Miseria in 2011, and Mexico is implementing a variant called Contigo Vamos Por Mas. Each program places different emphasis on the different components of the program: namely, demand-side factors and psychosocial support versus coordination between demand and supply.

8. The take-up of other social programs is not evaluated since the proportion of households using these programs at baseline is extremely low and there is insufficient statistical power.

9. See, for example, Alzúa, Cruces, and Ripani (2013).

hypotheses being tested simultaneously, we would expect to find some significant effects merely by chance. We therefore conclude that *Juntos* had no impact on the outcomes of interest overall. These results are also consistent with a preliminary evaluation of *Juntos* on a set of restricted labor market outcomes and other broader indicators, which found no impacts of the policy.[10]

Our main hypothesis for explaining why we observe no consistent impacts of the program is that treatment intensity was extremely low. Under the initial plan, social workers were intended to have an average caseload of 120 households per year under the intensive treatment arm and 180 households under a nonintensive (or classic) arm, and it was expected that households receiving the intensive treatment would experience the greatest positive effects. In practice, there was no distinction between intensive and nonintensive treatment: social workers received large caseloads, treating an average of 180 families per year across both arms. This has potential implications for both the quality and quantity of treatment. For example, households received an average of only three visits over an eighteen-month period across both treatment arms, which is much lower than the intended number of visits per year set out initially in the program plan. It is therefore unlikely that such treatment would have significant effects on household outcomes, even if such effects may be possible under a more intensive treatment scheme.

The home visits had two main objectives: to strengthen the psychosocial capabilities of the extreme poor that may be constraining their behavior, such as self-control; and to improve access to and use of available social programs through the provision of information and preferential access. It is unlikely that the first objective was reached with such a low number of visits. Furthermore, administrative data suggest that the way that *Juntos* operated implied a high degree of variation in the quality of home visits. In many municipalities, new social workers were hired each year to support targeted families, resulting in a lack of continuity in the relationship between the social worker and the household. Additionally, the qualifications and experience of social workers varied markedly.

The second objective constitutes an important channel through which targeted households could improve their economic outcomes if the social

10. A preliminary simple analysis of a restricted set of outcomes shows no consistent program impacts. This analysis was conducted in a short period to provide a quick assessment of the program impact on broad variables such as employment, income, and poverty at the household level, without considering gender differences.

programs were geared to their needs and were of sufficient quality. This channel could potentially be activated through the provision of information in the initial visits. However, knowledge about these programs and the use thereof do not seem to have improved significantly as a result of the intervention. Focus groups carried out by the initial evaluation consortium found that targeted households did not feel that they had preferential access. Moreover, they felt that a range of barriers prevented their access to these programs, including a mismatch between the program design (along many dimensions) and the needs of the target households. Consequently, if the supply of these programs is not improved, access and use will remain low.

Given the complexity of the program and the number of agencies involved, there might have been significant issues early on in terms of coordination and implementation that resulted in teething problems in setting up and running the program, but these problems should have dissipated over time as the program was rolled out nationally under the name *Unidos.* However, administrative data show that the treatment in the national program, *Unidos,* remains very light, despite being more intensive than in the pilot. Social workers in each municipality are assigned approximately 130 households per year, on average.[11] This contrasts with the case of *Chile Solidario,* which targeted a comparable population and formed the basis for the design of *Juntos* and later *Unidos.* Treatment in *Chile Solidario* was more intense, with 50 households per social worker on average—a much smaller caseload than in *Unidos.* As a result, households received an average of ten visits per year for a maximum of twenty-four months in *Chile Solidario.* In addition, households were guaranteed access to monetary subsidies to compensate them for participating in the program. Carneiro and others use a quasi-experimental approach to evaluate the effects of *Chile Solidario.*[12] They find a positive impact on the take-up of a family allowance for poor children (*subsidio único familiar*), but no impact for labor market or other economic outcomes.

Taken together with our results, this evidence has important policy implications not only for Colombia, but also for a wider context. *Unidos* is unlikely to make a significant contribution to the reduction of extreme poverty in Colombia. The results from the evaluation of *Juntos* suggest that the intervention had very little impact on the economic outcomes of participants in the short term and no impact on the take-up of existing programs. The evidence from *Chile Solidario* suggests that even a stronger version of this program

11. Information provided via private correspondence with ANSPE in August 2014.
12. Carneiro, Galasso, and Ginja (2014).

is unlikely to have significant impacts in improving the outcomes of the target population. These households are difficult to work with since they face constraints in different key areas such as skills, capital, and psychological traits. Although recent empirical evidence shows that some interventions are successful in alleviating such constraints in different developing countries, these are usually small-scale, high-quality interventions, often provided by a nongovernmental organization. A good recent example is a program to provide skills and vocational training to adolescent girls in Uganda.[13] In contrast, large-scale programs are usually provided through the welfare system, and there is a trade-off between quantity and quality. First, it may be extremely difficult to deliver high-quality interventions that provide good psychosocial support (through either home or group visits) on a large scale and at a reasonable cost, as the evidence discussed in this paper shows. Recent experimental evidence on how to use the infrastructure of *Familias en Acción* to deliver a scalable and integrated early childhood program through home visits in Colombia may provide some positive policy lessons in this area.[14] Second, even if the home-visit component of *Unidos* is effective, its impact is expected to be mediated through the use of other effective social programs. Hence, the effectiveness of a program such as *Unidos* depends on the quality of these other social programs and the extent to which they are tailored to the needs of the extreme poor.

Nevertheless, improvements in the program could potentially lead to more significant results. The necessary changes involve improving the quantity and quality of social workers, including the relationship or bond between the social worker and the households, and improving the supply of existing social programs in terms of quality and quantity. Properly investigating the impacts of an improved program would require conducting a further pilot program with an experimental evaluation. An experimental design could be used to determine whether improvements in the quality of social workers (for example, through better training or higher wages) and the reduction in caseload (through hiring additional social workers) lead to improvements in the policy impacts. It could also test whether some of the social programs available to the extreme poor are effective at all. This would indicate whether the program can be modified to have significant impacts or whether it should be replaced in its entirety.

13. Bandiera and others (2012).
14. Attanasio and others (2014).

The remainder of the paper is structured as follows. The next section provides background information on the program. We then describe the evaluation design and address some issues related to implementation. Subsequent sections discuss the data, provide descriptive statistics, and present our empirical methodology and results. The final section concludes.

Background and Description of the Program

In 2009, the Colombian government launched *Juntos,* a small-scale pilot of the *Unidos* program that would be rolled out nationally in late 2011. This is a social protection program for individuals living in extreme poverty in Colombia.[15] *Juntos* comprised the same objectives and design as the full-scale *Unidos* program, and participants faced the same eligibility criteria. *Unidos* is a government intervention that targets the population who live in extreme poverty, and currently serves nearly 1.5 million families in all 1,102 municipalities across the 32 departments of Colombia, at an annual cost of approximately US$140 million, or 5 percent of the total budget to promote social inclusion.[16] The scale and cost of the program clearly reflect its importance in Colombia.

The eligible population comprises two groups. First, households are eligible based on their low overall economic well-being. In Colombia, all households are categorized as one of six levels in the system for identifying potential beneficiaries of social subsidies (*Sistema de Identificación de Potenciales Beneficiarios de Subsidios Sociales,* SISBEN). SISBEN summarizes economic well-being and is used to identify eligible households for a number of different national welfare programs.[17] All households registered as SISBEN level 1 are eligible to enroll in the *Unidos* program, which includes roughly 20 percent of the poorest households. Around 1.2 million households qualify for *Unidos* under these criteria.

Second, households registered in the Central Registry for the Displaced Population (*Registro Único de Población Desplazada,* RUPD) are eligible

15. See the official website for more details (www.dnp.gov.co/Programas/DesarrolloSocial/Pol percentC3 percentADticasSocialesTransversales/RedUnidosparaSuperaci percentC3 percent B3ndelaPobrezaExtrema.aspx, last accessed on 26 February 2014).

16. Information provided by ANSPE via personal correspondence in August 2014, in turn sourced from "Reporte CIIF,"—Ministry of Finance and Public Credit and the Planning Advisory Office, Bogotá.

17. For more information, see www.sisben.gov.co.

to participate in *Unidos.* Colombia is among the countries with the highest proportion of internally displaced people in the world.[18] To be entered in the RUPD, households must prove that they have been internally displaced by providing an oral account of the facts to a public office. This population is considered to be largely marginalized from society, and the program aims to facilitate their use of existing social security programs.[19] Eligibility for displaced families is irrespective of their SISBEN classification, with many of these households classified at higher levels.[20] There are 300,000 such households targeted by the program.

These are the same criteria used for *Familias en Acción,* a conditional cash transfer program that targets poor households with children and whose positive impacts have been widely reported.[21] As a result, a significant proportion of households targeted by *Unidos* are already enrolled in *Familias en Acción.*

Unidos employs a two-pronged strategy to lift the most impoverished members of society out of poverty. The first prong aims to improve household skills and increase their demand for social programs through home visits and the provision of information on the programs, while the second strengthens the supply of existing social programs. The first prong has two specific objectives. The first objective is to improve people's knowledge of existing social welfare programs and facilitate their access to these programs by removing the constraints that prevent the poorest families from becoming recipients. For example, social workers can provide assistance in navigating the complex and confusing application processes for enrolling in existing programs. The second objective is to provide a sustainable long-term escape from poverty by helping families manage their own development, with a focus on specific strategic areas. This is expected to be achieved via the home visits, during which social workers (*cogestores social*) work with the families to identify

18. The United Nations Refugee Agency. See, for example, www.unhcr.org/pages/49c3646c23.html, last accessed 11 November 2014.

19. See Unidad para La Atención y Reparación Integral a las Víctimas (2013).

20. See section 2.4 of the 2009 *Juntos* operations manual.

21. As of June 2012, the law governing *Familias en Acción* also includes indigenous families (www.dps.gov.co/documentos/FA/LEY-FAMILIAS-ACCION.pdf). *Familias en Acción* is a countrywide conditional cash transfer program reaching 2.6 million families. It is aimed at encouraging beneficial health- and education-related behavior among deprived or displaced families with at least one child under the age of 18. Numerous papers assess the impact of this program. For example, Attanasio, Fitzsimmons, and Gómez (2005) report a positive impact on school enrollment and time spent in school; Baez and Camacho (2011) provide a useful and broad review of the evidence.

areas of vulnerability and to develop customized strategies or action plans, fitted to the unique circumstances of each family and taking into account their own capabilities, in order to address the identified issues and identify social programs that can help them to overcome these challenges. These strategies focus on nine key dimensions for sustainable development: personal identification cards; income and jobs; education and training; health; nutrition; housing; family dynamics; banking and savings; and access to justice.[22]

Social workers play an important role in achieving these objectives. The program is designed in theory to provide an intensive period of social support to households through home visits by social workers. These visits occur for up to five years, with the frequency of visits decreasing over time. The visits are divided into two stages. In the initial visits, the social worker works with the household to identify weaknesses and issues that they need to address in order to escape poverty. This is achieved through the completion of the family baseline questionnaire, which provides an assessment of 45 indicators (or *logros*). In the second phase, the household identifies the actions they need to take to achieve their objectives and, with the help of the social worker, devises a family plan that sets out the main priorities and how to address them. Follow-up visits are then made to the family, to provide support as needed and to check on progress toward the defined objectives. Households graduate from *Unidos* if they achieve all of their objectives within five years of enrolling in the program. According to the National Agency for Overcoming Extreme Poverty (ANSPE), around 16 percent of the beneficiary families had graduated from *Unidos* as of August 2014.[23]

The second arm of the program aims to improve access to existing social programs from the supply side. This is achieved in two ways. First, the program provides *Unidos*-eligible families with preferential access to existing social programs. Second, it aims to strengthen support for the agencies that manage the provision of welfare benefits. This is done to ensure that a sufficient supply of social welfare programs is available for all eligible households and that the programs meet the needs of the targeted population. The combination of

22. Our study focuses on the second dimension, income and jobs, since improvements in this area are most likely to raise households out of poverty in a sustainable way. Furthermore, the preliminary analysis on the impact of *Juntos* performed by the evaluation consortium showed no impact on any outcomes associated with the other dimensions. See Fedesarrollo, Econometría Consultores, SEI, and IFS (2012).

23. ANSPE (2013a, 2013b). See the official website for more details (www.dnp.gov.co/Programas/DesarrolloSocial/Pol percentC3 percentADticasSocialesTransversales/RedUnidospara Superaci percentC3 percentB3ndelaPobrezaExtrema.aspx, last accessed on 26 February 2014).

the two program arms should therefore serve the dual objectives of increasing the demand for social welfare programs among the poorest households, while simultaneously ensuring that a sufficient supply of these services is available to meet any increased demand.

The Evaluation of *Juntos*

The *Juntos* pilot program and its evaluation design were initial components of the wider *Unidos* program. The evaluation was planned in collaboration with ANSPE, the implementing government agency. This evaluation took place in seventy-seven municipalities, which were selected to provide a representative sample of all municipalities in Colombia.[24] As a result, our estimates should be interpreted as externally valid with respect to the impacts of the program across Colombia.

The evaluation employed an experimental design to ensure that individuals in treatment and control groups were comparable along observable and unobservable dimensions. Random assignment to treatment and control groups followed a structured process. First, the population of eligible families within participating municipalities was identified in early 2008. Second, each participating municipality was divided into several neighborhoods, or *barrios.* Third, between September 2008 and April 2009, each neighborhood was randomly assigned to one of four groups or cohorts. The program was rolled out to cohorts sequentially, so that the treatment began at different times across different neighborhoods. Given random assignment to cohorts, the characteristics of households across neighborhoods should be identical, on average, prior to the rollout of the program. This provides us with an opportunity to use neighborhoods in the fourth cohort as a control group for neighborhoods in the first cohort.

Given that the intensity of treatment was heterogeneous within the treatment group, the evaluation was designed to allow for a more detailed analysis of treatment impacts. More specifically, it allows us to test whether the impacts of the treatment varied with the intensity of treatment, as measured by the number of home visits received by the household. This was achieved

24. The consortium, consisting of Fedesarrollo, Econometría Consultores, the Institute for Fiscal Studies (IFS), and Sistemas Especializados de Información (SEI), who conducted the initial program design and evaluation, found that the selected municipalities did not differ in their observable characteristics from excluded municipalities. See Fedesarrollo, Econometría Consultores, SEI, and IFS (2012).

by further dividing the treatment group into classic and intensive treatment groups, as follows. First, social workers were recruited and randomly allocated to neighborhoods. Second, these social workers were randomly assigned to providing classic or intensive treatment. This process meant that household allocation into the two treatment arms was also random. Social workers who provided intensive treatment were, in theory, assigned fewer cases. This lower caseload would allow the social worker to focus more closely on each household and to provide a greater number of visits. This was not implemented in practice, however, as we discuss in the next section.

This design was intended to produce three distinct groups of interest: the control group (fourth cohort), the classic treatment group (first cohort), and the intensive treatment group (first cohort). Given random assignment, the characteristics of households across groups should be identical in the absence of the program. The impact of each treatment type can therefore be estimated by comparing mean outcomes between each treatment group and the control group in the post-program period. The evaluation design also separately identified three subpopulations of interest: rural, urban, and displaced. Even within the population of the extreme poor, the impacts of the program are likely to be highly heterogeneous across the three populations.[25] All subsequent analysis therefore examines each population separately.

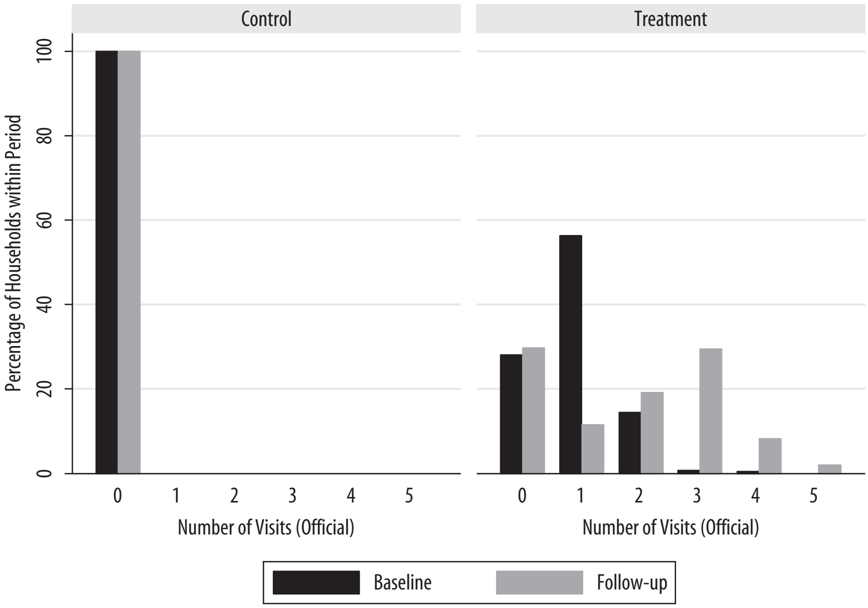
In this paper, we evaluate the short-run impact of *Juntos* from 2009 through mid-2011. The collection of baseline data occurred between November 2009 and March 2010. This period was prior to the initial treatment of all cohorts. Follow-up data were collected between June and August 2011, prior to the rollout of the program in neighborhoods assigned to cohort four. During the period between survey waves, households assigned to cohort one should have received home visits from social workers, while cohort four households should have received no visits. The evaluation finished in December 2011, and treatment was (in theory) subsequently rolled out to all eligible households across the country.

Evaluation and Program Implementation

Large-scale evaluations often face a number of challenges in their design and implementation. In this case, we need to consider two main issues when estimating the casual effect of the program. First, there is some contamination

25. Most of the displaced households are in urban areas (around 95 percent of the displaced households in our sample), so their behavior is likely to be most similar to urban households in a number of ways.

FIGURE 1. Number of Official Home Visits

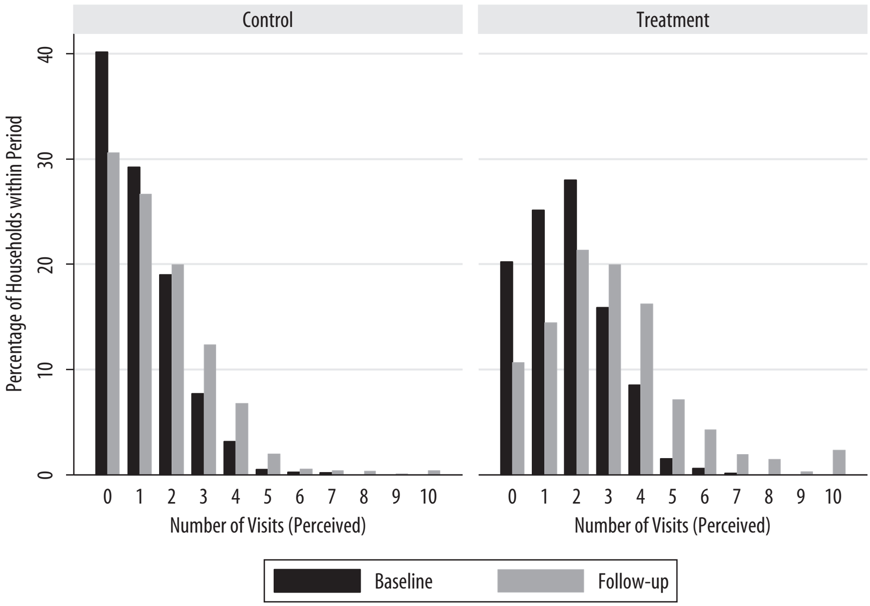


between the treatment and control groups. More broadly, a low number of visits are reported by all groups at follow-up. This suggests that treatment was only weakly implemented for the majority of participants. Second, households in the intensive treatment group did not systematically receive a higher number of visits than those in the classic treatment. We use two measures of the number of social worker home visits received by households in each wave to investigate these issues: the official number of visits recorded by the social workers; and the number of visits reported by the household in the household questionnaire (perceived visits).[26]

Figure 1 shows the number of official home visits made to households at baseline and follow-up and distinguishes between households assigned to the treatment and control groups. Figure 2 displays the same information for self-reported or perceived home visits. As explained above, social workers provide support to families through home visits, and these visits are, in principle, organized in sessions according to specific tasks. The social worker and the family coproduce the family baseline (first session of the home visits) and

26. We describe the surveys and questionnaires in more detail in the next section.

FIGURE 2. Number of Perceived Home Visits



family plan (second session of the home visits), and each session is expected to comprise two visits to the household. The number of sessions, and the associated number of visits, needed to implement the family plan is expected to vary across households according to their needs.[27] Together, these figures present three main points.

First, there is a large discrepancy between official and perceived visits. The initial evaluation suggests that some respondents may have mistaken the evaluation interviewer or public officials for social workers, and they therefore reported a higher number of visits than they received from the social worker. Anecdotal evidence also suggests that some social workers may have informally visited some households more frequently. Unfortunately, there is no information about social workers' characteristics that could shed light on potential variation in the quality of social workers and the quality of the visits in terms of their duration.

27. The implementation of the family plan mainly involves linking the families to the specific social programs that are tailored to their identified needs and strategic priorities organized around the nine dimensions discussed above. This information is from section 4.1.4 of the *Juntos* operations manual (24 March 2009).

Second, treatment was either weak or not administered at all for many members of the treatment group. Figure 1 suggests that 25 percent of the treatment group had received no visits at follow-up, while an additional 20 percent had received only one or two visits. Figure 2 shows a similar pattern in self-reported visits. This low intensity of treatment is unlikely to have produced significant changes in the outcomes of households over the period (even if a more intense version of the treatment would do so).

Finally, 70 percent of households in the control group reported at least one visit at follow-up. This is in contrast to the official data, which record no visits to households in the control group. This suggests that households in the control group were visited (perhaps informally) by a social worker or that they mistook an official or interviewer for a social worker.

Taken together, these figures suggest that some control group households received visits during the pilot phase. Meanwhile, many households who were assigned to the treatment group received no treatment or very weak treatment. As a result, randomly assigned treatment status may not accurately represent the actual treatment received.

Tables 1 and 2 show the average number of home visits at baseline and follow-up by treatment group, based on official and self-reported visits, respectively. These statistics are presented separately for each population.

TABLE 1. Average Number of Official Visits, by Treatment Group[a]

	Type of treatment		
Sample population	*Control*	*Classic*	*Intensive*
Displaced			
Baseline	0	0.63	0.60
	(0)	(0.69)	(0.69)
Follow-up	0	1.82	1.87
	(0)	(1.45)	(1.56)
Urban			
Baseline	0	0.95	0.98
	(0)	(0.68)	(0.58)
Follow-up	0	1.63	1.84
	(0)	(1.37)	(1.49)
Rural			
Baseline	0	1.08	1.04
	(0)	(0.68)	(0.64)
Follow-up	0	1.85	1.80
	(0)	(1.30)	(1.41)

a. The results in this table apply to the entire household panel collected in this survey. However, the results here closely reflect those obtained when examining only the subsample we use for our analysis.

TABLE 2. Average Number of Perceived Home Visits, by Treatment Group[a]

	Type of treatment		
Sample population	*Control*	*Classic*	*Intensive*
Displaced			
Baseline	1.00	1.60	1.42
	(1.09)	(1.21)	(1.24)
Follow-up	1.60	2.86	2.48
	(1.57)	(2.00)	(1.88)
Urban			
Baseline	0.96	1.84	1.76
	(1.18)	(1.37)	(1.44)
Follow-up	1.49	3.27	3.08
	(1.58)	(2.14)	(2.19)
Rural			
Baseline	1.21	1.91	1.81
	(1.20)	(1.31)	(1.39)
Follow-up	1.51	3.05	3.01
	(1.55)	(2.14)	(2.28)

a. The results in this table apply to the entire household panel collected in this survey. However, the results here closely reflect those obtained when examining only the subsample we use for our analysis.

The tables suggest that households in the intensive and classic treatment groups did not receive a significantly different number of (perceived or official) home visits at follow-up. Both treatment groups had on average approximately 1.8 official visits at follow-up, regardless of population type. The numbers of perceived visits were higher for all groups. Classic treatment households reported a higher number of visits at follow-up, on average, than intensive treatment households, but the difference is statistically insignificant.

These findings have two implications for our analysis. First, the absence of differences in the number of home visits between the classic and intensive treatment groups suggests that the analysis should ignore the distinction between the groups. We therefore group all treated households together in the remainder of this paper.

Second, given the issue of contamination, we obtain estimates using an instrumental variables (IV) approach, in addition to intention to treat (ITT) estimates. Specifically, we define two treatment dummy variables. The first is based on the assigned treatment (we call this variable T), which gives the ITT estimates. This variable takes the value of one if a household was originally allocated to intensive or classic treatment and zero otherwise, regardless of the number of official or perceived visits it actually received. The second

TABLE 3. Real versus Assigned Treatment[a]
Percent

Sample population	*Perceived control*	*Perceived treatment*
Displaced		
Assigned control	78.14	21.86
Assigned treatment	52.04	47.96
Urban		
Assigned control	78.03	21.97
Assigned treatment	40.42	59.58
Rural		
Assigned control	75.59	24.41
Assigned treatment	45.30	54.70

a. The results in this table apply to the entire household panel collected in this survey. However, the results here closely reflect those obtained when examining only the subsample we use for our analysis. The rows in the table indicate percentages within assigned groups.

dummy variable is based on self-reported or perceived visits by a *Juntos* social worker (we call this variable *real treatment,* or *RT*) at the time of the follow-up data collection. We consider this to be the real treatment, because it seems likely that only households that perceive visits from a social worker will be affected by the program, by increasing their knowledge and use of the available programs and thereby overcoming their extreme poverty. The variable *RT* takes the value of one if a household self-reported having received three home visits prior to the follow-up and zero otherwise. Hence, households that self-reported fewer than three visits at the time of follow-up make up the control group. We instrument this real treatment variable, *RT,* using the variable *T,* which reflects initial random allocation to treatment. We discuss the assumption underlying this empirical strategy in more detail below.[28]

Table 3 cross tabulates our real treatment (according to perceived visits) and randomly assigned treatment variables (*RT* and *T*) and summarizes the issue of contamination and imperfect compliance for the whole sample. As we explain in the next section, in our analysis we use a selected sample of households and individual members of these households, for which we observe a range of variables of interest in both periods. The patterns observed in figures 1 and 2 and tables 1 to 3 are very similar.

28. Our results are robust to a number of definitions. Results are robust if the official number of visits are used instead. Results change little if treatment is defined as two visits or four visits. These results are available on request.

Data and Descriptive Statistics

A rich set of data was collected as part of the evaluation. Data were collected in two separate waves. Initial data collection took place between November 2009 and March 2010, prior to the implementation of the *Juntos* pilot in cohort one neighborhoods (baseline). A second wave of data was collected between June and August 2011 (follow-up). The data contain a rich set of information, including sociodemographic characteristics at both the household and individual levels and individual labor market experiences. In addition, the follow-up data contain information on the knowledge and use of existing social welfare programs.

These data allow us to focus our analysis on three types of outcomes of particular importance given the aims of the program: namely, the knowledge of a range of existing social programs; the take-up of the main Colombian conditional cash transfer program, *Familias en Acción;* and labor market outcomes, in particular participation rate, employment rate (and type of employment), unemployment rate, hours worked, employment earnings, and tenure.

Given the scale of the evaluation, it was not feasible to sample the entire population of participants for the study. We use a random sample of participants collected across each of the municipalities. These samples were stratified by population type (urban, rural, and displaced) and the type of treatment (control, classic, and intensive). The sample size for each population was determined prior to data collection by power analysis. In addition, questionnaires of different lengths were administered to different households. Three types of questionnaire were administered (short, medium, and long) within cells defined by population type and the survey wave. These assignments were made according to power calculations specific to each outcome of interest.[29] Consequently, information is available for specific variables of interest for a subsample of households in both waves of the data. This means that we focus on the sample of households and individual members who provided a full set of information for all our variables of interest at both baseline and follow-up.

Program knowledge and usage information is contained only at follow-up in all cases. A total of 5,872 households were surveyed at baseline but we cannot use all of them in our analysis. Some households drop from the original sample due to classic attrition. Additional households were added at follow-up

29. The use of different questionnaires was due to a limited budget for data collection. For more details of this process, see See Fedesarrollo, Econometría Consultores, SEI, and IFS (2012).

to increase sample size, increasing the sample to 8,091. When we focus on households providing information in both waves, we have a balanced panel of 5,166 households.[30]

We further restrict our sample of interest to households with an adult head of household (aged 18 years or older), who provided a full set of answers to questions relating to labor market outcomes. This yields a final sample of 2,446 households. Our analysis also includes individual labor supply outcomes for individuals aged between eighteen and sixty years old. The final sample includes 5,042 individuals who fulfill these criteria.

To summarize, the selection of households and individuals into our sample may occur in three ways: classic household attrition (that is, households appear at baseline but are not included in the follow-up sample.); individual attrition (individuals appear at baseline but are not surveyed at follow-up, which may occur even if other members of their household remain in the sample); and questionnaire-type attrition (in principle, households were randomly assigned to different questionnaire types).[31]

One potential implication of the sample restrictions is that selection into the treatment and control groups is no longer random. We assess whether selection into the final sample was systematically related to treatment status in the following way. First, we take the entire sample at baseline and create a binary variable that takes the value of one if a household (or an individual) is in the final sample selected, and zero otherwise. Second, we regress the probability of appearing in the final sample on an indicator of whether the household was originally assigned to treatment, which is available to all households and individuals at baseline as opposed to other key variables for our analysis. We run a second regression that includes some baseline characteristics that are available for all observations and interacts these characteristics with an indicator of assigned treatment status. This tests whether the interaction of assigned treatment and baseline characteristics is systematically associated with appearing in the final sample among all households present at baseline. If this association is significant, the impact estimates obtained from the sample could be biased. For example, if the households that were initially assigned

30. Specifically, 13 percent of the initial sample did not appear in the follow-up survey, which is in line with the average attrition rates in large randomized controlled trials.

31. With regard to individual attrition, individuals may have left the household between waves. In addition, some individuals did not have consistent identifiers across the two periods. We match these individuals across waves using name and gender, successfully matching 82 percent of all individuals that appear in households in the panel.

TABLE 4. Impact of Assigned Treatment on the Likelihood of Being in the Household Sample[a]

	Displaced		*Urban*		*Rural*	
Variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Assigned treatment	−0.040	−0.328	−0.023	−0.111	−0.015	−0.091
	(0.051)	(0.214)	(0.024)	(0.153)	(0.022)	(0.160)
Baseline characteristics	No	Yes	No	Yes	No	Yes
Summary statistic						
F test[b]		1.816		2.175		1.161
P value		0.064		0.021		0.321
No. observations	1,872		1,720		2,280	

a. Columns 2, 4, and 6 control for pretreatment characteristics: age and education level of the household head, an indicator for whether the household head is also the household respondent; household size and composition; an index variable for municipal well-being; and the interaction of each of these with the treatment dummy. Clustered standard errors are reported in parentheses.

b. The *F* test (test of joint significance of interaction of demographics and treatment) tests the null hypothesis that the coefficients on all of the pretreatment characteristics interacted with treatment are jointly equal to zero.

to random treatment only remain in the sample if they have greater income than the households that leave the sample, we would overestimate the impact of the treatment on incomes. This analysis is conducted for each of the three populations to examine selection issues in each sample. We conduct a similar analysis at the individual level, by gender and by population type.

Table 4 shows the relationship between assignment to treatment and the likelihood of a household's appearing in the final sample using the sample of almost 6,000 households at baseline (note that only 2,446 households end up in our final sample). Results are displayed separately for each population type. The table provides two main insights. First, columns 1, 3, and 5 show that treatment status does not predict selection into the final sample across the three samples. However, columns 2, 4, and 6 show a slightly different picture. The *F* test, which tests the joint significance of all interactions between household baseline characteristics and assigned treatment status, is significant at the 5 and 10 percent levels for the urban and displaced populations, respectively. This suggests that selection into the final sample appears to be nonrandom for the urban and, to a lesser extent, the displaced population. In contrast, the results indicate that selection into the final rural sample is random.

Table 5 conducts a similar analysis at the individual level, considering over 14,000 individuals that appear at baseline, and shows the relationship between assigned treatment and the likelihood of an individual's appearing in the final sample of 5,042 individuals, by gender and population type. The results are similar to the household analysis, and indicate that the samples of male and female rural individuals are randomly selected. The male urban

TABLE 5. Impact of Assigned Treatment on the Likelihood of Being in the Individual Sample[a]

	Displaced				*Urban*				*Rural*			
	Female		*Male*		*Female*		*Male*		*Female*		*Male*	
Variable	*(1a)*	*(1b)*	*(2a)*	*(2b)*	*(3a)*	*(3b)*	*(4a)*	*(4b)*	*(5a)*	*(5b)*	*(6a)*	*(6b)*
Assigned treatment	−0.027	−0.213	−0.035	−0.392*	−0.024	−0.031	−0.020	0.143	0.013	0.015	−0.008	−0.052
	(0.049)	(0.229)	(0.048)	(0.212)	(0.026)	(0.124)	(0.027)	(0.127)	(0.023)	(0.137)	(0.025)	(0.144)
Baseline characteristics	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Summary statistic												
F test[b]		3.496		2.097		2.242		0.795		0.640		0.262
P value		0.000		0.029		0.017		0.634		0.778		0.988
No. observations	2,486		2,030		2,422		2,179		2,325		2,646	

*Statistically significant at the 10 percent level.

**Statistically significant at the 5 percent level.

***Statistically significant at the 1 percent level.

a. Columns numbered with *b* control for pretreatment characteristics: labor market status (active/inactive), age, education level, household size and composition, an indicator for the household respondent and household head, an index variable for municipal well-being, and the interaction of each of these with the treatment dummy. Clustered standard errors are reported in parentheses.

b. The *F* test (test of joint significance) tests the null hypothesis that the coefficients on all of the pretreatment characteristics interacted with treatment are jointly equal to zero.

sample also appears to be randomly selected, whereas the urban female and displaced individual samples remain nonrandomly selected.

Taken together, these findings suggest that the estimates of the program effects for the urban and displaced populations should be interpreted with some caution. To address this issue, we estimate regressions using first differences when data are available at baseline.[32] This accounts for permanent differences across individuals that could influence selection into the sample. This would reduce potential bias arising from the nonrandom selection into the sample for affected samples. We discuss our empirical strategy in more detail below.

Pre- and Post-Treatment Characteristics

This section presents descriptive statistics on the pre- and post-treatment income-generating activities of household heads and their key socio-demographic characteristics for our panel of households. In addition, we document the knowledge and use of public programs. We do not distinguish between the randomly assigned treatment groups. The following section explores differences across participants in this dimension.

We report unconditional means throughout this section. Therefore, statistics relating to employment, earnings, tenure, and hours all include zeros for those not active or unemployed. Consequently, observed changes in average earnings over time may be a result of either a genuine increase in earnings for individuals who are employed or an increase in the proportion of people who are employed and thus report any positive earnings.[33]

SOCIODEMOGRAPHIC CHARACTERISTICS AND LABOR MARKET OUTCOMES OF HEADS OF HOUSEHOLD. Table 6 shows the mean sociodemographic characteristics and labor market outcomes for our panel of households at both baseline and follow-up, by population type. Labor market outcomes refer to the heads of household in the panel sample. The table presents two interesting and broadly positive patterns in the labor market outcomes of these households. First, labor market participation remained relatively stable across the period. This is true for all three populations, with approximately 70 percent of household heads recorded as economically active.[34]

32. This is the case when examining labor market outcomes. Data on social program knowledge and use are unavailable at baseline, necessitating the comparison of levels at follow-up only.

33. Appendix A includes a detailed description of how the various variables were constructed.

34. See appendix A for the exact definition of *active* used.

TABLE 6. Basic Descriptive Statistics of Pre- and Post-Treatment Variables at the Household Level, by Population (Unconditional Means)[a]

	Displaced		Urban		Rural	
Descriptive statistic	*Baseline (1)*	*Follow-up (2)*	*Baseline (3)*	*Follow-up (4)*	*Baseline (5)*	*Follow-up (6)*
Labor market outcomes						
Active	0.72	0.73	0.70	0.72	0.68	0.68
	(0.45)	(0.44)	(0.46)	(0.45)	(0.47)	(0.47)
Employed	0.52	0.65	0.56	0.68	0.45	0.65
	(0.50)	(0.48)	(0.50)	(0.47)	(0.50)	(0.48)
Self-employed	0.21	0.33	0.26	0.36	0.18	0.32
	(0.41)	(0.47)	(0.44)	(0.48)	(0.38)	(0.47)
Wage earner	0.31	0.32	0.30	0.32	0.27	0.32
	(0.46)	(0.47)	(0.46)	(0.47)	(0.44)	(0.47)
Wage earner, formal	0.24	0.27	0.26	0.29	0.26	0.31
	(0.43)	(0.44)	(0.44)	(0.45)	(0.44)	(0.46)
Wage earner, informal	0.06	0.05	0.04	0.04	0.00	0.02
	(0.24)	(0.23)	(0.19)	(0.19)	(0.07)	(0.13)
Unemployed	0.20	0.09	0.14	0.04	0.24	0.03
	(0.40)	(0.28)	(0.35)	(0.18)	(0.43)	(0.17)
Wage and salary earnings	96,710.77	109,413.55	89,923.74	117,865.00	61,307.79	84,691.82
	(180,689.63)	(194,376.30)	(176,192.21)	(215,291.82)	(126,852.94)	(156,234.61)
Self-employment earnings	80,256.46	101,522.70	72,174.68	129,025.71	52,622.36	91,278.83
	(615,045.74)	(192,962.95)	(150,718.21)	(248,960.13)	(295,143.48)	(337,786.21)
Tenure	37.33	64.23	76.02	109.38	95.18	153.22
	(87.71)	(114.94)	(132.25)	(156.55)	(170.88)	(197.08)

Demographic characteristics						
Age	42.86	44.75	50.60	51.92	53.03	54.57
	(13.15)	(13.10)	(13.95)	(13.71)	(14.90)	(14.68)
Household respondent	0.66	0.65	0.59	0.56	0.60	0.55
	(0.47)	(0.48)	(0.49)	(0.50)	(0.49)	(0.50)
In relationship	0.66	0.66	0.73	0.70	0.73	0.73
	(0.47)	(0.48)	(0.44)	(0.46)	(0.44)	(0.45)
No. household members	5.23	5.08	5.45	5.00	4.80	4.86
	(2.26)	(2.15)	(2.51)	(2.39)	(2.41)	(2.42)
Male	0.50	0.51	0.64	0.69	0.75	0.77
	(0.50)	(0.50)	(0.48)	(0.46)	(0.43)	(0.42)
No. household members under 10	1.36	1.19	1.24	1.00	1.01	0.98
	(1.23)	(1.16)	(1.31)	(1.20)	(1.26)	(1.24)
No. household members over 60	0.24	0.27	0.45	0.46	0.54	0.58
	(0.53)	(0.57)	(0.68)	(0.70)	(0.74)	(0.77)
Years of schooling	4.59	4.82	3.61	3.68	2.33	2.30
	(3.68)	(3.82)	(3.31)	(3.39)	(2.55)	(2.55)
Municipal-level characteristics						
Municipality composite index	62.07	63.41	56.24	60.64	54.83	56.86
	(12.49)	(14.13)	(17.78)	(19.00)	(14.93)	(13.94)
No. observations	1,121	1,121	656	656	669	66,900

a. All of the labor market outcomes are unconditional variables in the sense that they take a value of one if true and zero otherwise. For example, *employed* equals one if the person in question is employed and zero if unemployed or inactive.

However, the composition of activity changed substantially over the period. Employment rates of household heads increased significantly within each population type, while unemployment fell substantially. For example, 52 percent of displaced household heads were employed in the baseline. This had increased to 65 percent for the same sample of households by the follow-up, an increase of 25 percent. Much of this growth in employment was driven by increases in self-employment. We also observe similar patterns for the other populations. The striking increase in employment rates and decrease in unemployment rates could be related to seasonality or to a sustained improvement in the labor outcomes of the extreme poor. These changes are clearly not a consequence of the introduction of the *Juntos* program, however, since we find no systematic and significant difference between treatment and control heads of households as discussed in the next section. A further investigation of the factors driving these changes is an important and interesting question for future research.

Second, wage and salary earnings and self-employment earnings of household heads increased over the same period. For example, in the displaced sample, the average head of household's wage and salary earnings at baseline was 96,710 Colombian pesos (COL$) per month (approximately US$50). This increased by 13 percent to COL$109,414 at follow-up. Even greater rises are observed in the other populations, with incomes growing by 31 percent and 38 percent for urban and rural households, respectively. Self-employment earnings increase proportionally more over time across the three populations, although the levels are lower than wage and salary earnings at baseline. These increases in employment income are largely driven by increases in the employment rate of the head of household.[35] Tenure also increased over the period, suggesting that employment was more sustainable.

Table 6 also reveals large differences in the sociodemographic characteristics of households across the three population types. These differences persist throughout the period. Displaced households have, on average, younger heads of household (forty-five years old at follow-up) relative to urban (fifty-two)

35. Despite these large increases in earnings, the monthly employment-conditional earnings or wages of these households' members remain below the monthly minimum wage, as expected for extreme poor households registered as SISBEN level 1. For example, take the individuals showing the highest employment-conditional earnings in our samples: male employees. Their average earnings conditional on employment were COL$393,616 (or approximately U$S198), $384,617 (US$193), and $261,732 (US$132) at follow-up for the displaced, urban, and rural samples respectively. The monthly minimum wage stood at Colombian $535,600 (or approximately US$269) in 2011.

or rural households (fifty-five). These heads of household are also more likely to be female and more likely to be the main respondent in the survey, less likely to be in a relationship, and have a higher level of education. There are no significant differences in the size of households across population type, although displaced households tend to be younger.

In the final row, we present a municipality-level composite index that reflects the quality of public service delivery in each municipality. This increases over time, and is relatively higher for displaced households. This indicates that displaced households are typically located in areas with a higher quality of public services. In contrast, rural households live in areas where the quality is lower.

Taken together, these characteristics suggest that displaced households generally live in better overall economic conditions than households in the other populations. Given that displaced households are eligible for enrollment in the program regardless of their SISBEN rating, evidence of such a pattern is not surprising.

KNOWLEDGE, USE, AND SUPPLY OF PUBLIC PROGRAMS. Table 7 presents the self-reported knowledge and usage of a selected group of public programs. These aim to provide support to individuals or households in order to foster income-generating activities. The table features programs that provide access to credit for micro enterprises, credits for education, and subsidized training activities. The table also includes the important program *Familias en Acción.* As mentioned in the previous section, *Familias en Acción* was launched in 2002 and is an established conditional cash transfer program aimed at improving the health and education outcomes of children in poor households.

The table shows the proportion of households who live in municipalities in which each program is active.[36] We do not condition on the availability of services in the municipality and so do not directly account for differences in the local supply of programs. Consequently, there is significant variation across populations and across specific programs in terms of their availability. Some programs are available in all municipalities. These include *Familias en Acción, Jóvenes Rurales Emprendedores* (which fosters income generating activities in rural areas), *Red Banca de las Oportunidades* (which provides access to formal microcredit, saving groups, and financial education for deprived households and individuals, as well as micro and small enterprises), and *Programa para el Desarrollo de las Oportunidades de Inversión y Capitalización* (which fosters income-generating activities of poor rural

36. This information was provided by the Unified Registry of Affiliates of the Social Protection System (RUAF) administered by the Colombian Ministry of Social Protection.

TABLE 7. Supply and Self-Reported Use and Knowledge of Public Programs Post-Treatment at the Household Level, by Population (Unconditional Means)[a]

	Displaced			*Urban*			*Rural*		
Program	*Supply*	*Knowledge*	*Use*	*Supply*	*Knowledge*	*Use*	*Supply*	*Knowledge*	*Use*
Familias en Acción	1.00	0.97	0.76	1.00	0.92	0.62	1.00	0.90	0.57
	(0.00)	(0.18)	(0.43)	(0.00)	(0.28)	(0.49)	(0.00)	(0.30)	(0.50)
Jóvenes en Acción	0.65	0.14	0.00	0.42	0.15	0.00	0.21	0.10	0.00
	(0.48)	(0.35)	(0.07)	(0.50)	(0.36)	(0.06)	(0.40)	(0.30)	(0.00)
Jóvenes Rurales Emprendedores	1.00	0.08	0.00	1.00	0.06	0.00	0.98	0.08	0.01
	(0.00)	(0.27)	(0.05)	(0.06)	(0.24)	(0.06)	(0.15)	(0.28)	(0.08)
Crédito ACCES del ICETEX	0.53	0.12	0.01	0.30	0.10	0.01	0.05	0.07	0.00
	(0.50)	(0.33)	(0.07)	(0.46)	(0.29)	(0.08)	(0.22)	(0.26)	(0.00)
Red Banca de las Oportunidades	1.00	0.11	0.00	1.00	0.09	0.00	1.00	0.07	0.00
	(0.00)	(0.32)	(0.05)	(0.00)	(0.28)	(0.04)	(0.00)	(0.26)	(0.07)
Generación de Ingresos de Acción Social	0.63	0.31	0.06	0.51	0.14	0.01	0.58	0.14	0.01
	(0.48)	(0.46)	(0.23)	(0.50)	(0.35)	(0.07)	(0.49)	(0.35)	(0.10)
Alianzas Productivas	0.20	0.07	0.00	0.20	0.06	0.00	0.25	0.08	0.00
	(0.40)	(0.25)	(0.04)	(0.40)	(0.24)	(0.00)	(0.43)	(0.28)	(0.04)
Programa para el Desarrollo	1.00	0.02	0.00	1.00	0.03	0.00	1.00	0.02	0.00
	(0.00)	(0.15)	(0.03)	(0.00)	(0.16)	(0.00)	(0.00)	(0.14)	(0.00)
Asistencia Técnica Rural	0.22	0.05	0.00	0.23	0.04	0.00	0.43	0.09	0.00
	(0.42)	(0.21)	(0.03)	(0.42)	(0.19)	(0.00)	(0.50)	(0.29)	(0.04)
No. observations		1,121			656			669	

a. As discussed in the main text, several programs appear to be used by almost none of the sample under consideration, and in many cases knowledge of these programs is very low.

households).[37] Other programs are unavailable to some households in our sample. For example, *Jóvenes en Acción,* which aims to provide vocational training for disadvantaged youth, is not available in all municipalities. This program was available to 65 percent of displaced households, but only 20 percent of rural households.

The table also shows that knowledge and use of the majority of the programs is very low. With the exception of *Familias en Acción,* the proportion of households who have knowledge of these programs ranges from zero to 0.15. Furthermore, use is extremely low, in most cases close to zero. Even programs specifically targeting the rural population are unknown and infrequently used by rural households in our sample.[38]

Overall, these findings suggest that knowledge and use of existing social programs is low among sample households. This highlights the importance of promoting and improving access to these programs for this population. However, it is concerning that knowledge and use is so low, particularly given that these statistics are reported after the intervention was launched.

Baseline Comparisons of Treatment and Control

The availability of baseline data allows us to test whether the randomly assigned treatment and control samples were balanced before the program started. If randomization was successful, baseline characteristics of those assigned to the treatment group (cohort one) will not differ in a statistically significant way from those assigned to the control group (cohort four). We test for balance in each household sample, based on household and head-of-household characteristics. We also test for balance in each individual sample for both genders.

HOUSEHOLD SAMPLES. Table 8 compares the baseline means of the assigned treatment and assigned control group for an array of household demographic characteristics and labor market outcomes, for each of the three populations.

37. For more information on *Jóvenes Rurales Emprendedores,* see www.sena.edu.co/oportunidades/emprendimiento-y-empresarismo/Jovenes%20Rurales%20Emprendedores/Paginas/Jovenes-Rurales-Emprendedores.aspx. For more information on *Red Banca de las Oportunidades,* see www.bancadelasoportunidades.com/contenido/contenido.aspx?catID=298&conID=673. For more information on *Programa para el Desarrollo de las Oportunidades de Inversión y Capitalización,* see www.minagricultura.gov.co/tramites-servicios/desarrollo-rural/Paginas/v1/Programa-desarrollo-de-las-oportunidades-de-inversion-y-capitalizacion-de-los-activos-de-las-microempresas-rurales.aspx.

38. Carneiro, Galasso, and Ginja (2014) report similar findings for the Chilean population enrolled in the *Chile Solidario* program.

TABLE 8. Baseline Differences between Treatment and Control Groups at the Household Level[a]

	Displaced		Urban		Rural	
Variable	*Control (1)*	*Treatment–control (2)*	*Control (3)*	*Treatment–control (4)*	*Control (5)*	*Treatment–control (6)*
Labor market outcomes						
Active	0.72	0.00	0.67	0.04	0.69	–0.01
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
Employed	0.51	0.01	0.56	0.00	0.49	–0.06
	(0.04)	(0.04)	(0.04)	(0.05)	(0.03)	(0.05)
Self-employed	0.20	0.02	0.26	0.01	0.20	–0.03
	(0.03)	(0.04)	(0.04)	(0.05)	(0.03)	(0.04)
Wage earner	0.31	0.00	0.31	–0.01	0.29	–0.04
	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
Wage earner, formal	0.05	0.02	0.03	0.01	0.01	–0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Wage earner, informal	0.26	–0.02	0.27	–0.02	0.28	–0.03
	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
Unemployed	0.21	–0.02	0.11	0.05	0.20	0.05
	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.04)
Wage and salary earnings	93,291.46	5,172.80	96,539.73	–9,797.04	67,111.28	–8,844.05
	(14,839.69)	(18,361.41)	(11,800.52)	(15,017.70)	(11,690.35)	(13,895.82)
Self-employment earnings	60,888.00	29,301.02	68,720.30	5,115.30	46,339.84	9,574.04
	(10,862.09)	(29,922.13)	(11,764.81)	(15,074.85)	(9,314.25)	(20,206.92)
Tenure	39.60	–3.44	80.83	–7.13	105.82	–16.22
	(6.07)	(7.17)	(8.36)	(11.07)	(12.09)	(16.33)
Demographic characteristics						
Age	44.37	–2.286**	50.76	–0.23	52.37	1.00
	(0.96)	(1.14)	(1.12)	(1.27)	(1.21)	(1.48)
Household respondent	0.62	0.063*	0.56	0.04	0.59	0.03
	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)

In relationship	0.69	−0.046*	0.70	0.05	0.72	0.01
	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
Number of household members	5.18	0.08	5.77	−0.467*	5.08	−0.427*
	(0.10)	(0.15)	(0.20)	(0.24)	(0.21)	(0.25)
Male	0.55	−0.062*	0.65	−0.01	0.74	0.01
	(0.02)	(0.04)	(0.04)	(0.05)	(0.03)	(0.04)
No. of household members under 10	1.34	0.02	1.40	−0.239**	1.13	−0.18
	(0.11)	(0.12)	(0.08)	(0.10)	(0.09)	(0.12)
No. of household members over 60	0.30	−0.103**	0.46	0.00	0.52	0.03
	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.07)
Years of schooling	4.44	0.23	3.29	0.48	2.44	−0.17
	(0.23)	(0.29)	(0.36)	(0.41)	(0.23)	(0.26)
Municipality level characteristics						
Municipality composite index	62.81	−1.12	56.20	0.06	54.05	1.19
	(3.10)	(3.67)	(4.39)	(5.40)	(2.76)	(3.60)
Summary statistic						
F test[a]	F(127,15) = 1.556		F(145,15) = 1.29		F(117,15) = 0.93	
P value	0.095		0.214		0.529	
Clusters	128		146		118	
No. observations, by group	380	741	213	443	230	439
No. observations, full sample	1,121		656		669	

*Statistically significant at the 10 percent level.
**Statistically significant at the 5 percent level.
***Statistically significant at the 1 percent level.
a. The wage earner, formal wage earner, informal wage earner, and unemployed variables have been omitted from the *F* test due to perfect collinearity. Standard deviations are in parentheses.

TABLE 9. Baseline Differences between Treatment and Control Groups at the Individual Level, by Gender[a]

	Displaced		*Urban*		*Rural*	
Gender	*F test (1)*	*P value (2)*	*F test (3)*	*P value (4)*	*F test (5)*	*P value (6)*
Female	F(120,17) = 2.012	0.02	F(141,17) = 1.758	0.04	F(112,17) = 0.961	0.51
Male	F(112,17) = 0.816	0.67	F(135,17) = 2.036	0.01	F(109,17) = 0.723	0.77

a. The variables included in the *F* test are the same variables included in table 8.

Columns 1, 3, and 5 report the baseline means of control sample households for displaced, urban, and rural households, respectively. Columns 2, 4, and 6 report the estimated difference between treatment and control households.

Overall, the results suggest that the three samples are highly balanced. When we conduct a test of joint significance of the differences in all baseline characteristics, we cannot reject the hypothesis that the characteristics of households in the treatment and control groups are the same in the urban and rural samples. The *F* statistics (*p* values) are 1.29 (0.214) and 0.93 (0.529) for urban and rural, respectively. This is consistent with very few individual statistically significant differences in certain characteristics when examined separately. The displaced sample is marginally unbalanced due to some imbalances in a number of household-head characteristics and the age of the household members. This translates into an *F* statistic (*p* value) of 1.556 (0.095). However, labor market outcomes seem balanced in this sample.

INDIVIDUAL SAMPLES. Table 9 presents the results of a test of joint significance of the differences in all baseline characteristics between individuals assigned to the treatment and cohort groups for displaced, urban, and rural individuals, respectively. We present results separately for males and females.[39] The first row of first two columns shows that the sample of displaced female individuals is not balanced. In particular, female individuals in the treatment group were more likely to be economically active and to be a formal wage earner (not displayed in the table). The *F* statistic (*p* value) for this sample is 2.012 (0.02). However, the sample of displaced male individuals (second row, first two columns of table 9) appears to be balanced, with an *F* statistic (*p* value) of 0.816 (0.67).

39. The coefficients for each individual characteristic for each sample are available on request. In some instances, we comment in the text on individual variables if these are driving some of the imbalances.

Columns 3 and 4 of the same table show that the samples of both males and females were not balanced at baseline for urban individuals. For females, this is driven by the fact that females in the treatment group were more likely to be formal wage earners, to live in smaller households, and to earn higher wages than their control-group counterparts. This results in an F statistic (p value) of 1.758 (0.04). Similarly, urban males in the treated sample lived in smaller households (with fewer children), although the labor market variables were not statistically different when tested separately. However, when tested jointly, the individual characteristics of males in the treatment group were statistically significantly different from males in the control group, with an F statistic (p value) of 2.036 (0.01). In contrast, columns 5 and 6 suggest that the individual samples of rural females and males were balanced. Despite some differences in characteristics when they were tested separately, we cannot reject the hypothesis that the characteristics of females and males in the treatment and control groups are the same when testing all variables jointly. The F statistics (p values) are 0.961 (0.51) and 0.723 (0.77) for females and males, respectively.

Together, the results suggest that only a subset of the individual samples is balanced: the rural samples and the displaced male sample. The individual samples are largely unbalanced for the displaced female individuals and both urban samples, driven to some extent by a few labor market outcomes. Given that the initial random assignment was made at the household level, the findings that some sample imbalances occur at the individual level are perhaps unsurprising. Many of the outcomes that we examine are at the household level, for which the samples appear to be balanced for all populations. Nevertheless, these results suggest that we should exercise some caution when interpreting the estimates of program impacts specifically on displaced and urban individuals.

Estimating Program Effects

Under a randomized controlled trial with no contamination between the assigned treatment ($T = 1$) and control ($T = 0$) groups, it is usually straightforward to identify the average treatment effect of a program by taking the difference in the empirical means of the outcome of interest between the two groups.[40] Since the evaluation under consideration has random assignment

40. Let $T = 1$ for those who were randomly assigned to treatment and 0 otherwise.

to treatment, this is the general approach that we adopt here to estimate the effects of the program. However, to control for the potential influence of contamination of the treatment and control group (that is, selection into the treatment group) and selection on observables and unobservables into our sample, we use the baseline information in our panel to augment this basic approach and ensure that our estimates are more robust. The precise approach we take is discussed in detail below.

Define y_i to be an outcome of interest for an individual (or a household) *i*. We can now write the expected average treatment effect (D) of *Juntos* on outcome y for the extremely poor households that received the treatment as follows: $D = E[y_i|T = 1] - E[y_i|T = 0]$. Since households were randomly assigned to treatment, we could obtain the average treatment effect by comparing the empirical means of the treatment and control group:[41]

$$\hat{D} = \hat{E}\left[y_i|T = 1\right] - \hat{E}\left[y_i|T = 0\right], \tag{1}$$

where $\hat{E}$ denotes the sample average. However, as documented earlier, there is substantial contamination between the randomly assigned treatment and control groups. Therefore, in the context of the *Juntos* evaluation under consideration, expression 1 reflects the intention-to-treat (ITT) estimate as opposed to the average treatment effect (ATE). To account for the fact that the evaluation design used cluster randomization, we can rewrite equation 1 in terms of a linear regression, where the cluster is a neighborhood denoted by *j*.[42]

This specification assumes that v_j and u_{ij} are i.i.d. with constant variance:

$$y_{ij} = \alpha + \beta T + v_j + u_{ij}. \tag{2}$$

41. See, for example, Duflo, Glennerster, and Kremer (2007) for an accessible review of impact evaluation methodologies.

42. Unfortunately, the survey data did not contain consistent identifiers of these neighborhoods. This variable is important for robust inference in the context of clustered samples. We therefore take a conservative approach and aggregate neighborhoods in bigger clusters defined by the three original different treatment levels (control, classic treatment, and intensive treatment) within each municipality. See, for instance, Pepper (2002) for a discussion of considering a more aggregate level of clusters in cluster samples. This gives a smaller number of clusters, hence decreasing the power of the analysis. As a robustness check, we calculate our impact results without clustering the standard errors, and the main results are consistent.

Under the stated assumptions, the estimated β reflects the impact of *Juntos* on the outcome of interest, y_{ij}. That is, $\hat{\beta}_{OLS} = \hat{E}[y_{ij}|T = 1] - \hat{E}[y_{ij}|T = 0] = 0$. While specification 2 is sufficient for the estimation of the effects of the *Juntos* program in theory, we take advantage of having panel data to increase the precision of our results and to ensure that they are robust. The approach we take is to augment specification (2) in three ways. First, to improve the precision of the estimates and control for any remaining baseline imbalances, which are important for urban individuals and displaced female individuals, we control for baseline (pretreatment) relevant characteristics $\mathbf{X}_{ik}$ (at the individual and municipality level, where municipality is denoted by k). This yields the following specification:

$$(3) \qquad y_{ijk} = \alpha + \beta T + \mathbf{X}_{ik}\gamma + \tilde{v}_j + \tilde{u}_{ijk}.$$

Second, as mentioned above, $\hat{\beta}_{OLS}$ in specification 3 will estimate the ITT estimate, but not the effect of the actual treatment on those that actually received visits by *Juntos* social workers. As described earlier, we define a variable that we call *real treatment* (RT), which takes the value of one if households received at least three (perceived) visits by the time of the follow-up data collection and zero otherwise. To estimate the effect of *Juntos* on those that actually received treatment, we need to use an instrumental variable approach. Therefore, we adopt the standard approach of using the assigned treatment variable (T) as an instrument for actual treatment (RT). By virtue of the fact that assignment to treatment was randomized, it should satisfy the standard independence and relevance assumptions:

$$(4) \qquad y_{ijk}\left|RT \perp T\right|\mathbf{X};$$

$$(5) \qquad \operatorname{cov}\left(RT, T\right) \neq 0.$$

In addition to the relevance assumption, we also need the stronger assumption of monotonicity—that is, the instrument makes every household either weakly more or less likely to actually participate in the *Juntos* program—which in this case is a reasonable assumption. Assignment to treatment should increase an individual's propensity to acquire treatment, and randomization should ensure that the exclusion restriction is satisfied, with assignment to treatment exerting no influence on the outcome variable, except through treatment itself. This provides identification of the treatment effect in the presence of contamination between the treatment and control groups:

$$\beta_{IV} = \frac{E\left[y_{ijk} \mid T_i = 1\right] - E\left[y_{ijk} \mid T_i = 0\right]}{P\left(RT_i = 1 \mid T_i = 1\right) - P\left(RT_i = 1 \mid T_i = 0\right)}. \tag{6}$$

Under the stated assumptions, this parameter is a measure of the average impact of the *Juntos* program on a particular outcome, y_i, for households (or individuals) in the sample that received the treatment (or the compliers) as a result of the random assignment.

Third, in our preferred specification, we take first differences of the outcome variables in order to remove any unobserved (time-invariant) differences in the level of the outcome variables that may have been present at baseline between the treatment and control group and that cannot be accounted for by observable characteristics. Removing unobserved time-invariant characteristics can also help correct for selection into our chosen sample that could generate a bias. In our empirical analysis, we report estimates using a difference-in-differences approach for both the ITT (ordinary least squares estimates using the assigned treatment) and the IV estimates.[43] The ITT specification is as follows:

$$\Delta y_{ijkt} = \gamma_0 + \beta T_i + \mathbf{X}_{ik,t-1} + \mu_{ijkt}. \tag{7}$$

To implement the IV approach in the difference-in-difference setup, we substitute T for RT in equation 7 and instrument RT with T as in the level regressions. Table 10 reports the first-stage regressions that predict the probability of a household's having reported that they received treatment, defined by the variable RT (which equals one if the household perceived having received at least three visits by a *Juntos* social worker by the time of the follow-up interview and zero otherwise), using assigned treatment as the instrumental variable, for each of the samples analyzed in this paper. The positive and significant coefficient on the assigned treatment variable indicates that, on

43. The results for the levels specification above (equation 3) are very similar. As mentioned above, the results are also robust to the precise cutoff used in defining the real treatment dummy variable and also to the use of perceived or official visits as a measurement of treatment. Results are available on request.

TABLE 10. First-Stage Regressions[a]

Explanatory variable	Displaced: Household head (1a)	Displaced: Individuals: Women (1b)	Displaced: Individuals: Men (1c)	Urban: Household head (2a)	Urban: Individuals: Women (2b)	Urban: Individuals: Men (2c)	Rural: Household head (3a)	Rural: Individuals: Women (3b)	Rural: Individuals: Men (3c)
Assigned treatment	0.295***	0.276***	0.221**	0.486***	0.507***	0.498***	0.430***	0.461***	0.522***
	(0.076)	(0.081)	(0.087)	(0.047)	(0.054)	(0.057)	(0.061)	(0.057)	(0.056)
Age	0.000	–0.001	–0.001	–0.001	0.000	0.001	0.000	0.000	0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Education	–0.006	–0.006*	–0.004	0.002	0.004	0.002	0.003	0.006	0.004
	(0.004)	(0.003)	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)
No. household members	–0.018*	–0.006	–0.003	0.014	0.005	–0.007	0.016	0.018	0.009
	(0.011)	(0.013)	(0.012)	(0.009)	(0.013)	(0.013)	(0.010)	(0.012)	(0.011)
No. household members under 10	0.046***	0.036*	0.036*	–0.022	–0.014	0.003	–0.021	–0.022	–0.011
	(0.017)	(0.020)	(0.020)	(0.021)	(0.025)	(0.026)	(0.020)	(0.023)	(0.020)
No. household members over 60	0.030	0.020	0.029	–0.073**	–0.044	–0.030	–0.041	–0.022	–0.050
	(0.031)	(0.040)	(0.040)	(0.033)	(0.040)	(0.039)	(0.035)	(0.038)	(0.050)
In relationship (= 1)	0.075*	0.058*	0.082*	–0.019	–0.002	0.059	0.034	–0.033	0.010
	(0.040)	(0.030)	(0.048)	(0.050)	(0.038)	(0.060)	(0.055)	(0.045)	(0.059)

(*continued*)

TABLE 10. First-Stage Regressions[a] (*Continued*)

Explanatory variable	*Displaced* Household head (1a)	*Displaced* Individuals: Women (1b)	*Displaced* Individuals: Men (1c)	*Urban* Household head (2a)	*Urban* Individuals: Women (2b)	*Urban* Individuals: Men (2c)	*Rural* Household head (3a)	*Rural* Individuals: Women (3b)	*Rural* Individuals: Men (3c)
Household respondent (= 1)	0.052	−0.009	0.012	−0.040	0.021	−0.060	−0.040	0.050	−0.054
	(0.049)	(0.041)	(0.053)	(0.037)	(0.032)	(0.047)	(0.041)	(0.045)	(0.046)
Municipality composite index	−0.005**	−0.005*	−0.004	−0.003**	−0.002**	−0.003*	−0.002	0.000	−0.001
	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Male (= 1)	−0.008			−0.067			−0.079		
	(0.039)			(0.046)			(0.048)		
Household head (= 1)		0.033	−0.075		0.036	−0.041		0.004	−0.030
		(0.036)	(0.050)		(0.040)	(0.051)		(0.060)	(0.063)
Constant	0.517***	0.522***	0.477***	0.433***	0.270**	0.309**	0.309**	0.088	0.105
	(0.176)	(0.184)	(0.175)	(0.139)	(0.128)	(0.137)	(0.157)	(0.131)	(0.150)
F statistic	15.000	11.500	6.470	106.910	89.040	75.450	50.180	65.710	87.850
P value	0.000	0.001	0.012	0.000	0.000	0.000	0.000	0.000	0.000
R squared	0.117	0.103	0.0776	0.229	0.243	0.245	0.175	0.196	0.253
Clusters	128	121	113	146	142	136	118	113	110
No. observations	1,121	1,354	966	656	790	648	669	632	652

*Statistically significant at the 10 percent level.

**Statistically significant at the 5 percent level.

***Statistically significant at the 1 percent level.

a. The dependent variable is 1(Perceived Treatment). The *F* statistic is equivalent to Kleibergen-Paap Wald *F* statistic and the Angrist-Pischke multivariate *F* test of excluded instruments. Robust and clustered standard errors are reported in parentheses.

average, this variable positively and significantly predicts having received treatment for all the samples. The *F* statistics to test for weak instruments are also reported; overall, these reject the hypothesis that the instrument is weak in each regression.

Knowledge and Use of Public Programs

We first look at the impact of *Juntos* on the knowledge and use of key social programs reported by the household survey respondent. One would expect that this would be one of the first areas in which a social worker would be able to have an influence, since the baseline knowledge of most social programs is low and providing knowledge of, and assisting these families in accessing, the programs they are eligible for seems like an appealing first step in helping to lift them out of poverty.

Table 11 shows the ITT and IV results for the level specification described in equation 3 only, since these variables were only collected at follow-up. As explained above, ideally we would like to estimate a difference-in-differences approach to deal with nonrandom selection into our panel sample, which is an issue particularly for urban households. Given data limitations, we cannot implement this approach for these outcomes and focus on the level specification. However, we can look at first differences for labor market outcomes as shown in the next section. We can only look at the effect of *Juntos* on the usage of *Familias en Acción,* since the usage of the other programs is close to zero (as shown in table 7) and there is insufficient variation across treatment status. Results are consistent with an increase in the knowledge of *Jóvenes Rurales Emprendedores,* significant at the 5 percent level, and in the knowledge of *Programa para el Desarrollo,* significant at the 10 percent level, by displaced households as a consequence of *Juntos* (columns 1a and 1b), although we have already seen that the sample of displaced households seems to suffer marginally from imbalances at baseline. Columns 6a and 6b show a positive impact on the proportion of rural households that use *Familias en Acción,* which is significant at the 10 percent level only. An ITT estimate shows that *Juntos* induced an increase of 7.5 percentage points in the probability of using *Familias en Acción:* the IV estimate is higher, at 17.3 percentage points. The overall take-up of *Familias en Acción* among the poor in both treatment and control rural households is estimated at 57 percent in our sample. Overall, given the large number of hypotheses being tested and the small number of statistically significant coefficients, only at the 5 or 10 percent level, we conclude that there were no positive impacts on the knowledge and use of social programs as a consequence of *Juntos.*

TABLE 11. Treatment Effect of *Juntos* on Knowledge and Use of Social Programs[a]

	Displaced				Urban				Rural			
	Knowledge		Use		Knowledge		Use		Knowledge		Use	
Program	ITT (1a)	IV (1b)	ITT (2a)	IV (2b)	ITT (3a)	IV (3b)	ITT (4a)	IV (4b)	ITT (5a)	IV (5b)	ITT (6a)	IV (6b)
Familias en Acción	0.009	0.031	0.046	0.155	0.007	0.015	0.015	0.030	0.037	0.087	0.075*	0.173*
	(0.013)	(0.042)	(0.039)	(0.134)	(0.041)	(0.084)	(0.044)	(0.089)	(0.033)	(0.079)	(0.043)	(0.104)
Jóvenes en Acción	0.014	0.048	—	—	0.006	0.013	—	—	0.021	0.049	—	—
	(0.045)	(0.158)			(0.035)	(0.072)			(0.029)	(0.065)		
Jóvenes Rurales Emprendedores	0.039**	0.130**	—	—	–0.031	–0.064	—	—	–0.003	–0.006	—	—
	(0.017)	(0.058)			(0.022)	(0.045)			(0.026)	(0.061)		
Crédito ACCES del ICETEX	0.000	0.099	—	—	–0.001	–0.029	—	—	0	–0.002	—	—
	(0.001)	(0.103)			(0.002)	(0.053)			(0.002)	(0.053)		
Red Banca de las Oportunidades	0.007	0.024	—	—	0.033	0.068	—	—	0.000	0.000	—	—
	(0.024)	(0.079)			(0.023)	(0.045)			(0.021)	(0.047)		
Generación de Ingresos de Acción Social	0.046	0.145	—	—	–0.01	–0.020	—	—	–0.043	–0.099	—	—
	(0.049)	(0.143)			(0.031)	(0.065)			(0.049)	(0.110)		
Alianzas Productivas	0.000	0.050	—	—	–0.001	–0.058	—	—	–0.000	0.004	—	—
	(0.001)	(0.052)			(0.002)	(0.043)			(0.002)	(0.058)		
Programa para el Desarrollo	0.013*	0.045	—	—	0.015	0.031	—	—	–0.004	–0.009	—	—
	(0.007)	(0.028)			(0.013)	(0.025)			(0.010)	(0.023)		
Asistencia Técnica Rural	0.004	0.013	—	—	–0.021	–0.043	—	—	–0.002	–0.004	—	—
	(0.013)	(0.043)			(0.017)	(0.035)			(0.028)	(0.063)		
No. observations	1,121				656				669			

*Statistically significant at the 10 percent level.

**Statistically significant at the 5 percent level.

a. OLS regressions estimate the ITT using assigned treatment. IV regressions instrument perceived treatment with assigned treatment. Regressions include the same baseline characteristics as those included in the first-stage regressions. The impact on program use cannot be estimated due to lack of variation for all programs except *Familias en Acción*. As discussed in the text, use is close to zero for most programs. Robust and clustered standard errors are reported in parentheses.

Labor Market Variables

Table 12 reports the ITT impact estimates and IV regressions for the first-difference specification described in equation 7 for the main outcomes of interest. First, overall the magnitude of the IV coefficients is larger than the magnitude of the ITT coefficients, as one might expect from the fact that the ITT coefficients use the assigned treatment variable to estimate the effect of treatment, and that there has been contamination of households assigned to the control group and imperfect compliance of those assigned to treatment.

Second, the IV standard errors are larger, again as one might expect. This is also true for the impact estimates for the other urban and rural populations discussed below. The only statistically significant result that holds across both the ITT and IV specifications is a positive impact of the *Juntos* program on the probability of being active for displaced household heads (column 1a). This result is robust to using the level specification described in equation 3. The IV results suggest that displaced heads of household are 27 percentage points more likely to be active as a result of the program than are heads of households that did not receive treatment. This is a substantial increase relative to the baseline level for households in the randomly assigned control group of 72 percent (as shown in table 8). This positive effect on the probability of being active for household heads is mirrored to some extent in the magnitudes of the estimates for the probability of being employed (column 2a) and the probability of being self-employed (column 3a). This may provide suggestive evidence that the impact on active status may be driven partially by the group that enters self-employment. However, these coefficients are not statistically significant.

The remainder of the impact estimates for the displaced sample suggest that there is no impact of the program on overall earnings, hours worked, or hourly pay. Given that the number of hypotheses being tested using only the IV specification in the tables for each of the three population groups (displaced, urban, rural) is thirty and that the number of significant coefficients is at maximum three, it could well be that these results are found by chance. Furthermore, the sample of displaced households is marginally unbalanced between treatment and control groups (as shown in table 8). Taken together, we interpret these results as indicative that the *Juntos* program did not have any effect on the labor outcomes of individuals in the displaced sample.

TABLE 12. Treatment Effect of *Juntos* on Participation, Employment, and Earnings, Displaced Population, First-Difference Specification[a]

Estimation method	Active: Household head (1a)	Active: Individuals: Women (1b)	Active: Individuals: Men (1c)	Employed: Household head (2a)	Employed: Individuals: Women (2b)	Employed: Individuals: Men (2c)
ITT impact estimation	0.081**	0.015	0.048*	0.056	0.005	0.019
	(0.034)	(0.034)	(0.027)	(0.060)	(0.051)	(0.071)
IV regressions	0.273**	0.053	0.215	0.191	0.019	0.085
	(0.122)	(0.118)	(0.144)	(0.183)	(0.181)	(0.307)
	Unemployed			*Hours worked per week*		
	Household head (5a)	*Individuals: Women (5b)*	*Individuals: Men (5c)*	*Household head (6a)*	*Individuals: Women (6b)*	*Individuals: Men (6c)*
ITT impact estimation	0.024	0.009	0.029	3.387	0.152	1.559
	(0.048)	(0.035)	(0.065)	(3.638)	(2.669)	(4.332)
IV regressions	0.082	0.034	0.130	11.465	0.550	7.041
	(0.172)	(0.132)	(0.310)	(11.130)	(9.495)	(18.733)
	Hourly wage			*Tenure*		
	Household head (9a)	*Individuals: Women (9b)*	*Individuals: Men (9c)*	*Household head (10a)*	*Individuals: Women (10b)*	*Individuals: Men (10c)*
ITT impact estimation	−177.6	16.7	−89.9	7.8	3.9	4.0
	(326.9)	(87.5)	(135.8)	(15.3)	(5.8)	(14.0)
IV regressions	−601.3	60.5	−406.3	26.4	14.1	18.2
	(1,108)	(305.4)	(684.5)	(48.3)	(18.7)	(59.6)
Clusters	128	121	113	128	121	113
No. observations	1,121	1,354	966	1,121	1,354	966

*Statistically significant at the 10 percent level.

**Statistically significant at the 5 percent level.

a. The ITT coefficients use the assigned treatment variable to estimate the effect of treatment. The IV regressions use perceived treatment, with assigned treatment as the instrumental variable. Regressions include the same baseline characteristics as those included in the first-stage regressions. Robust and clustered standard errors are reported in parentheses.

Table 13 shows similar results for the urban sample. From table 9, we know that individual samples showed imbalances between assigned treatment and control at baseline, and assigned treatment was shown to be systematically associated with the probability of being in the sample for the household-level sample (see table 4). Results should therefore be considered with some caution. Only the coefficient on the probability of being active for the sample of women is statistically significant in this table—and only at the 10 percent level. Overall, there seems to be no consistent impact of *Juntos* on the population of individuals living in urban areas.

	Self-employed			*Wage earner*		
	Household head (3a)	*Individuals: Women (3b)*	*Individuals: Men (3c)*	*Household head (4a)*	*Individuals: Women (4b)*	*Individuals: Men (4c)*
	0.079	−0.006	0.017	−0.023	0.011	0.002
	(0.054)	(0.034)	(0.064)	(0.049)	(0.033)	(0.057)
	0.269	−0.021	0.077	−0.077	0.040	0.008
	(0.175)	(0.125)	(0.279)	(0.172)	(0.112)	(0.253)
	Wage and salary earnings			*Self-employment earnings*		
	Household head (7a)	*Individuals: Women (7b)*	*Individuals: Men (7c)*	*Household head (8a)*	*Individuals: Women (8b)*	*Individuals: Men (8c)*
	−24,654	−8,641	−19,468	−28,265	200	−4,946
	(22,958)	(14,768)	(25,153)	(45,044)	(7,621)	(18,448)
	−83,444	−31,279	−87,943	−95,666	724	−22,341
	(86,864)	(57,610)	(123,258)	(154,468)	(27,275)	(86,114)
	128	121	113	128	121	113
	1,121	1,354	966	1,121	1,354	966

Table 14 reports the results for the rural sample. Here, three out of thirty coefficients are significant in the IV results, both for levels and first-difference specifications. This again suggests that the results could be found by chance. Furthermore, the only significant results indicate a negative impact for the sample of women of the *Juntos* program on the probability of being employed, mirrored by a decrease in the probability of being self-employed and a decrease in hourly wages or pay (largely due to a composition effect or a decrease in the number of individuals that are employed in the first place). However, this negative impact on female self-employment, when combined

TABLE 13. Treatment Effect of *Juntos* on Participation, Employment, and Earnings, Urban Population, First-Difference Specification[a]

	Active			*Employed*		
Estimation method	*Household head (1a)*	*Individuals: Women (1b)*	*Individuals: Men (1c)*	*Household head (2a)*	*Individuals: Women (2b)*	*Individuals: Men (2c)*
ITT impact estimation	−0.044	−0.082*	−0.015	−0.006	−0.050	−0.004
	(0.044)	(0.048)	(0.049)	(0.052)	(0.047)	(0.061)
IV regressions	−0.090	−0.163*	−0.030	−0.012	−0.099	−0.007
	(0.091)	(0.097)	(0.097)	(0.107)	(0.094)	(0.121)
	Unemployed			*Hours worked per week*		
	Household head (5a)	*Individuals: Women (5b)*	*Individuals: Men (5c)*	*Household head (6a)*	*Individuals: Women (6b)*	*Individuals: Men (6c)*
ITT impact estimation	−0.038	−0.032	−0.011	−0.208	−1.259	4.869
	(0.030)	(0.035)	(0.034)	(3.170)	(2.050)	(3.421)
IV regressions	−0.078	−0.064	−0.022	−0.428	−2.484	9.773
	(0.062)	(0.068)	(0.067)	(6.460)	(4.025)	(6.866)
	Hourly wage			*Tenure*		
	Household head (9a)	*Individuals: Women (9b)*	*Individuals: Men (9c)*	*Household head (10a)*	*Individuals: Women (10b)*	*Individuals: Men (10c)*
ITT impact estimation	89.4	4.3	−23.9	6.8	−5.0	10.9
	(155.1)	(130.7)	(171.3)	(17.8)	(6.1)	(11.8)
IV regressions	184.1	8.5	−48.0	14.0	−9.9	21.8
	(315.5)	(255.1)	(339.3)	(36.0)	(12.3)	(23.4)
Clusters	146	142	136	146	142	136
No. observations	656	790	648	656	790	648

*Statistically significant at the 10 percent level.

a. The ITT coefficients use the assigned treatment variable to estimate the effect of treatment. The IV regressions use perceived treatment, with assigned treatment as the instrumental variable. Regressions include the same baseline characteristics as those included in the first-stage regressions. Robust and clustered standard errors are reported in parentheses.

with the marginally significant positive impact on the take-up of *Familias en Acción,* is consistent with recent empirical evidence that looks at the impact of conditional cash transfer programs on labor supply and finds a small negative impact for rural women.[44]

44. See, for example, Alzúa, Cruces, and Ripani (2013), who provide evidence of a small negative effect on the probability of being employed for rural women of the conditional cash transfer program *Progresa* in Mexico.

Self-employed			Wage earner		
Household head (3a)	*Individuals: Women (3b)*	*Individuals: Men (3c)*	*Household head (4a)*	*Individuals: Women (4b)*	*Individuals: Men (4c)*
0.067	−0.000	−0.021	−0.073	−0.050	0.018
(0.043)	(0.033)	(0.042)	(0.047)	(0.035)	(0.060)
0.138	−0.001	−0.043	−0.15	−0.098	0.036
(0.087)	(0.065)	(0.083)	(0.098)	(0.067)	(0.119)
Wage and salary earnings			*Self-employment earnings*		
Household head (7a)	*Individuals: Women (7b)*	*Individuals: Men (7c)*	*Household head (8a)*	*Individuals: Women (8b)*	*Individuals: Men (8c)*
−13,193	−13,837	31,085	7,663	964	−14,999
(19,048)	(11,129)	(21,438)	(21,177)	(13,966)	(18,028)
−27,163	−27,295	62,401	15,777	1,901	−30,110
(38,961)	(21,940)	(42,644)	(43,279)	(27,233)	(35,264)
146	142	136	146	142	136
656	790	648	656	790	648

Discussion and Concluding Remarks

This paper provides an evaluation of the initial phase of the large-scale intervention *Unidos* (that is, the pilot program, *Juntos*) in relation to its impact on access to social programs and on the labor market outcomes of the extreme poor in Colombia. This paper also makes use of a rich data set to provide a detailed description of the labor market lives of this traditionally understudied population.

TABLE 14. Treatment Effect of *Juntos* on Participation, Employment, and Earnings, Rural Population, First-Difference Specification[a]

	Active			Employed		
Estimation method	*Household head (1a)*	*Individuals: Women (1b)*	*Individuals: Men (1c)*	*Household head (2a)*	*Individuals: Women (2b)*	*Individuals: Men (2c)*
ITT impact estimation	−0.011	−0.081	0.001	0.049	−0.084*	0.063
	(0.045)	(0.068)	(0.041)	(0.046)	(0.044)	(0.063)
IV regressions	−0.026	−0.175	0.001	0.114	−0.182*	0.121
	(0.104)	(0.149)	(0.078)	(0.107)	(0.099)	(0.118)
	Unemployed			*Hours worked per week*		
	Household head (5a)	*Individuals: Women (5b)*	*Individuals: Men (5c)*	*Household head (6a)*	*Individuals: Women (6b)*	*Individuals: Men (6c)*
ITT impact estimation	−0.060	0.003	−0.062	2.984	−1.720	4.059
	(0.040)	(0.051)	(0.053)	(2.350)	(1.774)	(2.990)
IV regressions	−0.140	0.007	−0.119	6.935	−3.728	7.778
	(0.095)	(0.110)	(0.097)	(5.397)	(3.859)	(5.546)
	Hourly wage			*Tenure*		
	Household head (9a)	*Individuals: Women (9b)*	*Individuals: Men (9c)*	*Household head (10a)*	*Individuals: Women (10b)*	*Individuals: Men (10c)*
ITT impact estimation	159.4	−200.3**	146.1	18.2	−9.2	23.3
	(196.6)	(79.9)	(134.2)	(20.1)	(9.0)	(20.0)
IV regressions	370.4	−434.2**	280.0	42.2	−20.0	44.5
	(449.3)	(188.8)	(250.2)	(45.9)	(19.7)	(37.1)
Clusters	118	113	110	118	113	110
No. observations	669	632	652	669	632	652

*Statistically significant at the 10 percent level.
**Statistically significant at the 5 percent level.
a. The ITT coefficients use the assigned treatment variable to estimate the effect of treatment. The IV regressions use perceived treatment, with assigned treatment as the instrumental variable. Regressions include the same baseline characteristics as those included in the first-stage regressions. Robust and clustered standard errors are reported in parentheses.

In terms of the evaluation, we find no consistent short-run impact of *Juntos* on our outcomes of interest: knowledge of a range of existing social programs; the take-up of the main Colombian conditional cash transfer program, *Familias en Acción;* and a range of labor market outcomes, including the participation rate, employment rate (and type of employment), unemployment rate, hours worked, employment earnings, and tenure. The estimated zero effect of the program on labor market outcomes is not surprising: the program failed to

Self-employed			*Wage earner*		
Household head (3a)	*Individuals: Women (3b)*	*Individuals: Men (3c)*	*Household head (4a)*	*Individuals: Women (4b)*	*Individuals: Men (4c)*
0.021	−0.074**	0.006	0.028	−0.010	0.057
(0.058)	(0.035)	(0.079)	(0.049)	(0.027)	(0.069)
0.049	−0.160**	0.011	0.065	−0.022	0.110
(0.135)	(0.081)	(0.150)	(0.111)	(0.058)	(0.127)
Wage and salary earnings			*Self-employment earnings*		
Household head (7a)	*Individuals: Women (7b)*	*Individuals: Men (7c)*	*Household head (8a)*	*Individuals: Women (8b)*	*Individuals: Men (8c)*
13,297	−4,808	18,945	4,423	−5,499	6,326
(14,410)	(5,367)	(18,503)	(31,2320)	(4,541)	(17,342)
30,906	−10,422	36,304	10,279	−11,922	12,122
(32,846)	(11,543)	(34,142)	(71,663)	(10,048)	(32,883)
118	113	110	118	113	110
669	632	652	669	632	652

influence participants' knowledge or use of the available social programs, and the availability of several public programs is low for both the treatment and control groups (see table 7). One would expect that if the *Juntos* intervention were to have a substantive effect, the knowledge and use of social programs would be one of the first important constraints that would be relaxed.

As discussed above, we believe that the lack of impact is largely driven by the lightness of the treatment (that is, the low number of home visits received by treated households). Given that *Unidos,* the national counterpart of the pilot *Juntos,* seems to have very similar features, the evidence would suggest that it is unlikely to transform the lives of the extreme poor in Colombia. Even if social workers have a lower average caseload under *Unidos* than under *Juntos,* this is still more than double the caseload that social workers have under the similar program *Chile Solidario.* Moreover, although *Chile Solidario* seems to have an effect on the take-up of social programs, it still does not significantly transform the lives of the poor in Chile. This could be rationalized by a lack of direct impact of home visits on the behavior of the poor beyond the take-up of social programs and by uncertainty around the quality of social programs available through the welfare system, the take-up of which is supposed to be incentivized through programs such as *Unidos* and *Chile Solidario.*

Furthermore, working with these households is particularly difficult given the multitude of constraints they face in different key areas. This includes a lack of skills and capital and the presence of certain psychological traits. Banerjee and others provide empirical evidence showing that the multifaceted Graduation program implemented in six different developing countries can generate progress in reducing extreme poverty, by improving self-employment income and well-being more generally.[45] This program uses a multipronged approach that sequentially tackles capital, skill, psychological constraints, and informational restrictions. The authors document that these programs are very costly to run, though their calculations suggest they are cost-effective in most countries. However, the scale of these programs is small, covering fewer than 11,000 households in six countries. In contrast, *Unidos* aims to cover around 1.5 million households at a national scale using only public resources. How to optimize and cost-effectively implement programs such as Graduation at a national scale in countries with limited state capacity remains an open question.

In this context, we reiterate the suggestion to policymakers that the *Unidos* program requires substantial reforms, and these reforms should be evaluated to understand what (if anything) is effective for which population and the mechanisms through which these impacts occur. Three specific areas should be addressed: an expansion of the supply of programs and widespread promotion to boost knowledge of their existence among potential beneficiaries; an assessment of which programs are most effective in improving the lives of the poor, in order to guide the selection of programs that should be made available

45. Banerjee and others (2015).

through *Unidos* and thereby improve not only the quantity but also the quality of social programs; and an increase in the budget allocated to social workers to ensure that they are suitably skilled and have workloads at least in line with other similar programs (such as *Chile Solidario*). Importantly, policymakers should strive to guarantee the quality of these evaluations, by avoiding the contamination of the samples and the dilution of the treatment. This last point seems particularly crucial in the case of this evaluation, in which a failure to fully implement the program due to a lack of coordination across the different state agencies involved, combined with insufficient funds, led to a poorly implemented policy in both the pilot phase evaluated in this paper and in its final, scaled-up version.

Given the evidence presented in this paper, it seems unlikely that in its current form, *Unidos* is having a substantial positive influence on the livelihoods of the extreme poor in Colombia. One would hope that by iteratively adjusting and improving the program and appropriately evaluating these changes to learn what works and what does not, we might converge on a more cost-effective way to assist this population. Precisely how to do this remains an open question left to future work.

Appendix: Data Description and Definitions

The data used in this paper were collected primarily for the evaluation of the *Juntos* program. A secondary advantage of the data set is that it contains extremely detailed descriptive and behavioral information pertaining to the lives of a very understudied group of individuals—namely, the poorest members of society. Data collection comprised two waves: a baseline survey, conducted between November 2009 and March 2010 before the start of treatment, and a follow-up survey, conducted between June 2011 and August 2011, after the treatment group began treatment.

The survey included seventy-seven municipalities, chosen to be representative of the country as a whole. Each municipality was divided into several neighborhoods (clusters), which served as the unit of randomization. Clusters were randomized into one of four cohort groups. Each of these groups commenced with treatment at a different point in time. The impact analysis in this paper compares the outcomes of cohort 1, which received treatment first and therefore was treated prior to the second wave of data collection, with cohort 4, which received treatment last and therefore were designated to be untreated at the second wave of data collection.

Survey Structure

The survey consisted of two parts. The first part collected information on the characteristics of the household, as well as general information on all the members of the household. This part consisted of several detailed modules relating to different aspects of the lives of these individuals. The module containing questions regarding knowledge and use of social programs is of particular interest to the current analysis in this paper. The second part of the survey collected detailed health, education, and labor market information at an individual level for all the relevant members of the household.

To satisfy the resource constraints of any project, there is often a trade-off made between the size of the sample and the level of detail of the survey. In this project, this trade-off was addressed by conducting a shorter survey to a wide sample, while administering a more detailed survey to a smaller subsample of individuals. Consequently, there were two types of questionnaires at baseline (long and short) and three types of questionnaires at follow-up (long, medium, and short). The short questionnaire contained core questions that were asked of every household, while the medium and long questionnaires asked individuals more detailed information and were administered to a subset of households. The allocation of households to each questionnaire type was done randomly and therefore should not have influenced the selection of our sample for analysis. This is examined in some detail in the main text.

Matching Individuals across Waves

Because of the way in which the data were encoded, individuals in the data set were not assigned a personal identifier number that corresponded across the two waves. Therefore, while it was straightforward to match households across waves, it was slightly more challenging to match the individuals within these households across waves. To do so, we used the names and birthdates of the individuals. However, there appeared to be a substantial number of inconsistencies in both of these variables.[46] We therefore employed a matching algorithm that used the available information to match individuals who lived

46. For example, there were frequently spelling mistakes in names, or first and second names were often switched. In addition, on examining the raw data, we often found that an individual who was clearly the same person in baseline and follow-up had a deviation in either the day, month, or year of birth between the two waves.

in the same household in both waves and appeared to be the same person, up to a small number of errors in their recorded data.

The matching algorithm started by matching individuals within a given household with date of birth and name data that agreed perfectly across waves, and thereafter we matched using a sequence of criteria that relaxed perfect consistency along each of these dimensions. At every step of this matching process, we matched only among individuals who are in the same household at baseline and follow-up, and we matched only among the unmatched individuals. Therefore, by starting with the strictest criterion for matching individuals and moving to more relaxed criteria, we limited the chance of making an incorrect match, as may occur if we were to only use the most relaxed matching criterion. The guiding principle behind this method for matching individuals was to strike a balance between matching as many individuals as possible and minimizing the likelihood of making an incorrect match.

The way we implemented this matching process was as follows. First, we matched only those individuals within the same household across waves who have exactly the same name and date of birth recorded in both waves, with no mistakes. Second, among the unmatched individuals, we allowed for small lexicographical deviations and common spelling mistakes, provided the date of birth is the same. Third, we relaxed perfect consistency along the date of birth dimension, by allowing for one deviation in either year, month, or day of birth, provided the individual's first and last name matched between waves. Fourth, we allowed for small errors along both dimensions. Fifth, among the remaining unmatched individuals, we matched individuals who had a perfect match for either first and last name or for their date of birth only. Finally, we manually checked the remaining unmatched individuals within households that were observed in both waves. After completing this process, we selected 1,000 individuals randomly to check for accuracy of the procedure; this exercise showed the procedure was extremely accurate. In the end, around 82 percent of the individual members of households appearing in our panel of households were matched.

Variable Definitions

Active status: the active variable is an indicator variable, defined for individuals over the age of seventeen. It takes a value of one if the person is either currently employed, has spent the majority of the last week working, or has searched for work in the last four weeks; it takes a value of zero otherwise.

Employment: the employed variable is an indicator variable that takes a value of one if the person has listed at least one job in which he or she is currently employed, and a zero otherwise. Since this variable takes a value of zero for inactive individuals, it reflects unconditional employment, as opposed to employment conditional on being active. We consider that this way of defining employment status makes much more sense than defining employed as being employed at any point in the last year in the context of the evaluation of the *Juntos* program. Since treatment only began during the year, it would be a very noisy measure to consider any employment during the preceding year; even in the follow-up questionnaire, much of this employment would have occurred prior to treatment.

Self-employed and wage earners: employed workers in our data set are divided into two categories: those who work for a wage and those who are self-employed. We therefore define a dummy variable called wage earner, which equals one for individuals who state that they are currently employed in a job in which they earn a wage, and zero otherwise. Correspondingly, the self-employed variable is a dummy variable equal to one for all individuals who state that they are either self-employed or a business owner, and zero otherwise. Both these variables take a value of zero if the individual is unemployed or inactive.

Formal and informal workers: wage earners are classified as either formal or informal workers on the basis of whether they reported holding an employment contract. The formal wage earner variable takes a value of one for workers who report holding a contract and zero otherwise. Similarly, the informal wage earner variable takes a value of one for workers who report not holding an employment contract.[47]

Wage and salary earnings and self-employed earnings: wage and salary earnings are the monthly wages or salaries reported by individuals whose primary current job was as a wage earner. Self-employment earnings are

47. In this sample, the group of self-employed workers would fall into the category of informal workers under most internationally used *informality* definitions. However, since this group is quite different from the set of informal wage earners, we examine the two groups separately. We also considered two alternative definitions of *informality*: first, one that defines a worker as informal if he or she works in a firm with fewer than six employees; and second, one that combines the two definitions, with workers defined as informal if they do not hold a contract or if they work in a firm with fewer than six employees. The contract informality and firm informality definitions are highly correlated and yield similar results. Once we use the combined contract and firm size definition, the proportion of the wage earners defined as informal increases to approximately 90 percent.

the monthly earnings net of costs for the self-employed. For those who are unemployed or inactive, we impute zero values. In addition, for those who reported earning a minimum wage, we did not observe their monthly salary or earnings. We therefore imputed these values using their monthly hours worked and the national minimum wage for the relevant point in time.

Hours worked and hourly earnings: for the set of individuals who reported currently holding a job, this variable reflects their self-reported number of hours worked in the last week. In order to calculate the hourly earnings, we multiply the weekly wage by 4.33 to get an approximate number of hours worked in the month. We then divide the wage and salary earnings of individuals who are wage earners, or self-employed earnings for self-employed individuals, by this monthly hours worked variable to obtain an hourly earnings.

Tenure: the tenure variable reports the number of months that the individual has spent in his or her current job, truncated at the date of the interview. This variable was calculated using the start date reported for employed individuals' current job. For unemployed or inactive individuals, we impute a zero value for the tenure variable.

Composite municipality level index: This variable reflects a municipality level variable that is a composite index reflecting the quality of public service delivery in each municipality.

References

Alzúa, María Laura, Guillermo Cruces, and Laura Ripani. 2013. "Welfare Programs and Labor Supply in Developing Countries: Experimental Evidence from Latin America." *Journal of Population Economics* 26(4): 1255–84.

ANSPE (Agencia Nacional para la Superación de la Pobreza Extrema). 2013a. "Familias UNIDOS y asistencia de los niños al colegio." Bogotá.

ANSPE. 2013b. "Las Familias UNIDOS en el 2012." Bogotá.

Attanasio, Orazio, Emla Fitzsimmons, and Ana Gómez. 2005. "The Impact of a Conditional Education Subsidy on School Enrollment in Colombia." Report Summary Familias 1. London: Institute for Fiscal Studies.

Attanasio, Orazio, and others. 2014. "Using the Infrastructure of a Conditional Cash Transfer Program to Deliver a Scalable Integrated Early Child Development Program in Colombia: Cluster Randomized Controlled Trial." *British Medical Journal* 349 (September).

Baez, Javier E., and Adriana Camacho. 2011. "Assessing the Long-Term Effects of Conditional Cash Transfers on Human Capital: Evidence from Colombia." Policy Research Working Paper 5681. Washington: World Bank.

Bandiera, O., and others. 2012. "Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda." Working Paper. London School of Economics.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101(2): 274–98.

Banerjee, Abhijit V., and others. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348(6236).

Carneiro, Pedro, Emanuela Galasso, and Rita Ginja. 2014. "Tackling Social Exclusion: Evidence from Chile." Discussion Paper 8209. Bonn: Institute for the Study of Labor (IZA).

Currie, Janet. 2006. "The Take-up of Social Benefits." In *Poverty, the Distribution of Income, and Public Policy,* edited by Alan J. Auerbach, David Card, and John M. Quigley, pp. 80–148, New York: Russell Sage.

Dalton, Patricio S., Sayantan Ghosal, and Anandi Mani. 2014. "Poverty and Aspirations Failure." *Economic Journal* (forthcoming).

Duflo, Esther. 2012. "Human Values and the Design of the Fight against Poverty." Tanner Lecture presented at Harvard University, May. University of Utah, Tanner Humanities Center.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics Research: A Toolkit." *Handbook of Development Economics,* vol. 4, edited by T. Paul Schultz and John A. Strauss, pp. 3895–962. Amsterdam: Elsevier.

Fedesarrollo, Econometría Consultores, SEI (Servicios Especializados de Información), and IFS (Institute for Fiscal Studies). 2012. "Evaluación de impacto de juntos (hoy unidos): red de protección social para la superación de la pobreza extrema informe de evaluación, diciembre de 2011." Technical Report. Bogotá.

Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60(1): 35–52.

Ghatak, Maitreesh, and Nien-Huei Jiang. 2002. "A Simple Model of Inequality, Occupational Choice, and Development." *Journal of Development Economics* 69(1): 205–26.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108(3): 551–75.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much.* New York: Macmillan.

Pepper, John V. 2002. "Robust Inferences from Random Clustered Samples: An Application Using Data from the Panel Study of Income Dynamics." *Economics Letters* 75(3): 341–45.

Unidad para La Atención y Reparación Integral a las Víctimas. 2013. "Informe Nacional de Desplazamiento Forzado en Colombia 1985 a 2012." Technical Report. Bogotá.