Online appendix for:
Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment[+]

Karen Macours, Norbert Schady, and Renos Vakis

Appendix 1: Design of *Atención a Crisis* and its evaluation

Further information on the *Atención a Crisis* interventions

From November 2005 until December 2006, the Ministry of the family in Nicaragua (MIFAMILIA) implemented the *Atención a Crisis* pilot program in 6 municipalities in the northern part of the country. The municipalities were selected for their extreme levels of poverty and because they had been affected by a severe drought in the previous year. The program had two objectives. First, it aimed to serve as a short-run safety net by providing cash transfers to reduce the need for adverse coping mechanisms, such as taking children out of school or reductions in food consumption. Second, the program intended to promote long run upward mobility and poverty reduction by enhancing households' asset base and income diversification capacity.

[+] We are grateful to Ximena Del Carpio, Fernando Galeana, and Patrick Premand for countless contributions during data collection and preparation, the program team at the Ministerio de la Familia and in particular Teresa Suazo for their collaboration during the design of the impact evaluation, and CIERUNIC, the Centro de Investigación de Estudios Rurales y Urbanos de Nicaragua (in particular Verónica Aguilera, Carold Herrera, Enoe Moncada, and the entire field team) for excellent data collection. This research benefited from World Bank support including the Trust Fund for Environmentally & Socially Sustainable Development (TFESSD) made available by the governments of Finland and Norway, the Bank-Netherlands Partnership Trust Fund Program (BNPP) as well as the Research Committee through a Research Support Budget (RSB) grant. Funding from BASIS was also received under the USAID Agreement No. EDH-A- 00-06-0003-00 awarded to the Assets and Market Access Collaborative Research Support Program. We thank Jere Behrman, Sally Grantham-McGreggor, Elizabeth King, John Maluccio, Christina Paxson, Elisabeth Sadoulet, Miguel Urquiola, seminar participants at the World Bank, the Inter-American Bank, FAO, the Institute for Fiscal Studies, Berkeley, Leuven, PSE, Paris 2, and Wisconsin-Madison, two anonymous referees and the editor Esther Duflo for providing us with very helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the World Bank or the Inter-American Development Bank, the Executive Directors of these organizations, the countries they represent, or any of its affiliated organizations. All errors and omissions are our own.
Contacts: karenmacours@parisschoolofeconomics.eu, norberts@iadb.org and rvakis@worldbank.org

A total of 3002 households were selected to participate in the program. These households were allocated one of three different packages through a participatory lottery:

(i) A conditional cash transfer (CCT).

(ii) A conditional cash transfer plus a scholarship that allowed one of the household members to participate in a vocational training course.

(iii) A conditional cash transfer plus a productive investment grant, aimed at encouraging recipients to start a small non-agricultural activity.

All selected beneficiary households received the basic CCT. All households, including those without children, received bi-monthly payments, adding up to a total transfer of US $ 145 during the year of the program. Households with children between 7 and 15 enrolled in and attending primary school received an additional US $ 90 per household, and an additional US $ 25 per child (with all amounts referring to the total transfer received over the year), conditional on school enrollment and attendance. The school enrollment and attendance requirement was carefully monitored by the ministry, through data received from the primary school teachers. The program was also meant to include a condition that required that households with children under the age of 6 would take these children to health centers for preventive health check-ups. However, due to implementation problems, these visits to the health centers were not monitored by the program (Aguilera et al. 2006).

In addition to the CCT, one third of the beneficiary households also received a scholarship that allowed one of the adult household members (preferably a member between 15 and 25 years of age) to choose among a number of vocational training courses offered in the municipal headquarters. In addition to covering the costs of the training, the program also compensated the participants for lost wages while in training, up to 6 months (US $15 per month). The scholarship was conditional on regular attendance at the course. The courses aimed at providing participants

with new skills for income diversification outside of subsistence farming. These beneficiaries were also offered labor-market and business-skill training workshops organized in their own communities.

Finally, another third of the beneficiary households received, in addition to the CCT, a grant for productive investments aimed at encouraging recipients to start a small non-agricultural business activity with the goal of asset creation and income diversification. This grant was conditional on the household developing a business development plan, outlining the objectives of the business and proposed investments in new livestock or non-agricultural income-generating activities. Beneficiaries received technical assistance to make a business plan and also participated in business-skills training workshops organized in their own communities.

Due to implementation delays, the vocational training courses had not started at the moment the data of the 2006 follow-up survey were collected. They took place in the fall of 2006. At the time of the 2006 survey, the difference between the vocational training beneficiaries and those of the basic CCT package was hence limited, though they might have had different expectations about future skills, about related future income and/or expectations about compensation for the time spent in training. The beneficiaries of the productive investment package, on the other hand, had received the largest amount of benefits: at the end of May 2006 they had received $175 to invest in a nonagricultural activity. The remaining $25 was to be paid on the next payment day (after survey completion). In addition they had received technical assistance to select the activity and develop a business plan, help which they were still receiving during the 2006 follow-up survey.

All beneficiaries of the *Atención a Crisis* program, regardless of the treatment they were assigned to, were exposed to repeated information and communication efforts by program staff during enrollment and pay-days. These stressed the importance of varied diets, health, and education, and were meant to change household investment and consumption patterns. Program

participants were also required to participate in a number of local events and talks ranging from discussions on nutrition practices to workshops on business development and labor market skills. A subset of beneficiary women were selected during the registration assemblies to serve as *promotoras* or leaders of small groups of beneficiary women (approximately 10 per group) in order to further enhance information flows and motivation and to encourage compliance with the various program requirements and conditionalities. Specifically, the *promotoras* were expected to frequently meet with the beneficiaries in their groups to talk about the objectives and the conditionalities of the program.

Further information on the impact evaluation design

The program was targeted to 6 municipalities in the Northwest of Nicaragua.[1] From the list of all communities in the 6 municipalities, 56 intervention and 50 control communities were randomly selected through a lottery to which the mayors of the 6 municipalities were invited to attend and participate.[2] Baseline data on household assets and household composition were then used to define program eligibility, resulting in the identification of 3002 households to participate in the program. This amounts to more than 90 percent of the households in treatment communities, and includes 95 percent of households with children under 6 years old. The eligibility criteria were

[1] The budget for the pilot allowed targeting 3000 households for a one-year period, which was much smaller than the population of the 6 municipalities. The program was therefore allocated randomly. Households were notified that funding of the project implied that the program would last 1 year, and would only cover the treatment communities. Households in the control communities did not receive any program benefits. They were notified that if there was a decision to scale up the program after the initial year, the control communities would be incorporated. People in the treatment communities understood the program was only to last for a year, and people in the control communities knew that there was a possibility they might receive the program the next year. They also knew it was likely to depend on the result of the national elections that were to be organized at the end of that year. In that election, the government changed and the project was not scaled up.

[2] Before the lottery, all communities in the 6 municipalities were grouped in pairs based on similarity in road access and microclimate. Through the lottery, one community of each pair was selected as a treatment community, the other as control. In case of uneven number of communities, a "pair" consisted of the largest community and the combination of the two other communities. The identification of communities and community pairs was based on maps and discussions with municipality technical personnel. Communities tend to be geographically separated from each other, which reduces the potential for spillover effects from the treatment to the control communities.

determined using the proxy means methodology developed for the RPS and based on the national household data from 2001 (EMNV). Additional discussions with local leaders from each intervention community were conducted to identify possible exclusion or inclusion errors. Based on the discussions with leaders, 3.7 percent of all the households considered were re-assigned from non-eligible to eligible, and 3.7 percent from eligible to non-eligible. To avoid any possible selection bias resulting from the re-assignment by the leaders, the results we present use eligibility by the proxy means as the intent-to-treat (without taking into account the reclassification by the community leaders).

From each eligible household, the female household member who was reported as the primary caregiver was then invited to a registration assembly.[3] During the assemblies, the program objectives and its various components were explained, women were asked to enroll in the program, and *promotoras* were selected on a voluntary basis. Among the intent-to-treat households, enrollment in the program was about 95%.[4] At the end of each assembly, all the beneficiaries participated in a lottery process through which the three packages described above were randomly allocated among the eligible households. Specifically, each beneficiary was asked to randomly draw a ball with 1 of 3 colors from a black, nontransparent bag. For each assembly, the bags contained an equal number of balls of each color, and the total number of balls matched the total number of beneficiaries in the assembly. At the end of the day each color was matched to an intervention package through another lottery attended by all beneficiaries from the community. Hence, at the moment households signed up for the program, they did not know which of the three interventions

[3] In the few cases where there was no adult female in the household, an adult male was selected as the program recipient.

[4] Of the 5 percent that did not enroll, the majority are households that had been reclassified by the leaders. The remainder are households that had migrated out of the treatment communities by the start of the program, and (very few) cases of households that refused the program.

package they would end up receiving.[5] Take-up of the CCT component was the same for all packages. Take up of the additional packages was also high. Among households enrolled in the program, 89 percent of the households eligible for the vocational training grant enrolled a household member in a course. Take-up of the productive investment grant among eligible households in the program was almost 100 percent. About 10 percent of the business development plans had initially been refused by the Ministry of the Family, but these were sent back to the households and virtually all of them developed a new plan, with the help of technical assistance (with the few exceptions being households that migrated out after the registration assemblies).

Given the objectives of the program, the outcomes that the impact evaluation as a whole focuses on are related to household investments in human capital and productive assets. Among human capital outcomes, there was a strong interest on early childhood development outcomes, the focus of this paper.[6] This was motivated by anecdotal evidence of early childhood delays in the region of study during baseline pretesting, and given the lack of evidence on CCT effects on ECD outcomes. With this objective, a cognitive test (TVIP) was added to the baseline survey, and data on anthropometrics were collected. In both follow-up surveys a more extensive set of tests was conducted (see appendix 3).

Data were collected at baseline on all households in the treatment communities, and for a random sample of households in the control group. The sample size in the control communities was chosen to be equal to one-third of the population in the treatment, in order to obtain a control group of equal size to each of the treatment groups. The proxy means test used to identify eligible households in the treatment communities was also used to identify households that would have been eligible in the control communities.

[5] Due to the transparency of the process, the lottery process was widely perceived as fair. Participation by the invited beneficiaries to the assemblies and lotteries was near 100%.
[6] See Macours and Vakis (2009) for short-term impacts on other outcome variables.

Appendix 2: Attrition

Attrition can potentially introduce serious biases into the estimation of program effects. In this study, attrition between the baseline and follow-up surveys was minimal. Only 1.3 percent and 2.4 percent of households interviewed at baseline could not be re-interviewed in 2006 and 2008, respectively. Among children less than 6 years old when the transfers started and living in the sample households at baseline, 4.7 percent of children could not be re-interviewed in the 2006 follow-up, and 3.1 percent in the 2008 follow-up. The low attrition rates are a result of repeat visits to recover temporary absence and extensive tracking of migrants. Migrant households and children were interviewed and tested in their new locations.

Attrition is uncorrelated with treatment—in a regression of attrited households on a dummy for treatment, the coefficient is -0.004 in 2006 (with a standard error of .005), and 0.003 in 2008 (with a standard error of 0.006). In a comparable regression for children the coefficient for 2006 is 0.005, with a standard error of 0.013, and for 2008 the coefficient is 0.0004 with a standard error of 0.009. Similar results are obtained when considering each of the treatment packages separately, and when comparing between treatment packages. Appendix Table A1 shows that the baseline characteristics of the full sample of households and those that could be located at follow-ups are very similar.

Aside from attrition because of failure to re-interview, 6 percent of the children did not do one or more of the tests that rely on their active participation (Denver sub-scores, TVIP, memory, leg motor) in 2006. This was typically due to refusal to participate by extremely shy children, who were not willing to interact with the test administrators in a way that allowed the test to be conducted. This is of potential concern if the treatment affected the willingness of the children to participate. Appendix Table A2 shows that the baseline characteristics of children who did all tests

are very similar to those of all children located at follow-up. The share of children who refused to take at least one test is lower in treated than in control communities in 2006, but the difference is not significant (the coefficient is 0.014 and the standard error 0.012). There are also no significant difference between the basic package and the lump sum payment package (coefficient -0.009, standard error 0.012).

In addition, there are some other children for whom the full set of outcome variables is not available, either because of missing anthropometric measures, which was typically due to logistical problems with the anthropometric material (6 percent of children in 2006) or because the behavioral problem index is missing, typically because the caregiver could not be located by the test administrator (8 percent of children in 2006). There is no significant difference between treatment and control in the likelihood that the full set of outcome variables is available (the coefficient is 0.014 and standard error 0.019).

In 2008, field teams did more repeat visits and attempts to convince the children to participate in case of refusals and to recover missing information for anthropometrics or the behavioral index. As a result, the share of children who refused to participate in one or more tests is only 4 percent, and there is no significant difference between treatment and control communities (the coefficient is -0.003, and the standard error 0.009), nor between the basic and the lump-sum package (coefficient -0.004, standard error 0.009). Similarly, there are fewer children for whom anthropometric measures are missing (1 percent), or for whom the BPI is missing (an additional 4 percent). Here too there are no significant difference between treatment and control on the likelihood that the full set of outcome variables is available or not (the coefficient is as 0.004 and standard error 0.014).

Given that the differences between treatment and control are small and not significant, and given that we control for baseline differences, any resulting selection bias is likely to be small. This

is further confirmed by a robustness check in which only the sample of children that have completed all the tests for their age group is used. The results from this robustness check are very similar to the main results we report—an average program effect of 0.08 standard deviations in 2006 (0.1 standard deviations for cognitive-social emotional outcomes, and 0.05 for health and motor development) and an average program effect of 0.07 standard deviations in 2008 (0.08 standard deviations for cognitive-social emotional outcomes and 0.07 for health and motor development).

Appendix 3: Measurement of early childhood development tests and intermediate inputs

Early childhood development tests

We focus on eleven measures of early childhood development. Height and weight were measured using standardized anthropometric material and procedures for all children. Height-for-age and weight-for-age were calculated using the international norms.

Social-personal, language, fine motor, and gross motor skills for children between 0 and 83 months old were assessed using the four sub-tests of the Denver Developmental Screening Test (Frankenberg and Dodds 1996). For each subtest, the child is asked to perform a number of age-specific tasks. When children fail to perform a task that 75 percent of children of their age in the reference population can perform, the test falls back to easier tasks, up to the point where tasks are reached that the child can perform.[7] In case certain behaviors or tasks cannot be observed, the caregiver is asked about the ability of the child to perform them. The social-personal subtest mainly consists of behavior that the caregiver is asked about, such as social interactions, the ability of a child to dress and eat on their own, imitate others, etc. The language subtest covers recognition and use of sounds, words, sentences, etc. The fine motor skills subtest mainly relates to observations of manual tasks such as drawing, playing with cubes, reaching for objects, etc. Finally, the gross motor tasks capture observations of basic crawling, sitting, walking, as well as throwing, jumping, etc. For the language, fine motor and gross motor subtests, most items are scored through direct observation of whether the child can perform the task. A relatively small share (15 out of 39 for language, 1 out of 29 for fine motor, and 5 out of 32 for gross motor) can be scored by asking the caregiver about the child's ability for that specific task, in case the item cannot be directly observed. In the sample, 40 (37) percent of language tests, 8 (4) percent of fine motor tests and 29 (23) percent of gross

[7] Similarly, for children performing all tasks for their age group the test continues with more difficult tasks. For the children in our sample this occurred very rarely.

motor tests have at least one item administered through the caregiver's report in 2006 (2008). Importantly, neither in 2006 nor in 2008 is there a significant difference between treatment and control on the likelihood that the Denver items were obtained uniquely through direct observation, or in the number of items obtained through self-reporting (P-values of the significance tests range from 0.23 to 0.97). Hence, caregivers in the treatment are not more likely to provide answers on behalf of their children, further reducing concerns regarding potential differences in reporting. Moreover, as discussed in section 4, our findings are robust to excluding the Denver subtests from the estimations.

The Denver scores are based on the number of tasks a child fails to perform, when these tasks can be carried out by more than 90 percent of children of the same age in the reference population.[8] The Denver is designed for children between 0 and 6 years of age. For this study, the test was also applied to somewhat older children, given the substantial delays in cognitive development that exist in our sample (described in detail in the paper).[9] For children age 36 months or older we applied five additional tests. The first of these is the TVIP, the Spanish-speaking version of the Peabody Picture Vocabulary Test (PPVT), a test of receptive vocabulary that has been widely used in developed and developing countries.[10] Children are shown a series of slides with four pictures each (for example, the first slide has a picture of a flashlight, a boat, a basket, and a hot-air balloon), and are asked to point at a given object stated by the enumerator (for example, "boat"). Test items gradually become more difficult. The test administrator records the number of correct and incorrect responses, and the test stops when a child is making as many errors as she

[8] The Denver has been used in other studies of early childhood development in developing countries, including in Nicaragua (Oberhelman et al. 1998). Other applications in developing countries include Halpern et al. (1996); Cheung et al. (2001); Choudhury and Gorman (2003); and Dewey et al (2001).

[9] Results are robust when only including the Denver test results for children up to 72 months of age (to reflect the age range the Denver was originally designed for). They are also robust to alternative ways of scoring the Denver test (accounting for the number of delays and cautions, or alternatively using a binary indicator on whether the child has at least one, or alternatively two, delays).

[10] See, for example, Paxson and Schady (2007, 2010), Schady (2011), Umbel et al. (1992), Baydar and Brooks-Gunn (1991), Blau and Grossberg (1992), Rosenzweig and Wolpin (1994), and Fernald, Gertler, and Neufeld (2008).

would be expected to make if she were randomly guessing.[11] We also use a short-term memory test and a leg motor test from the McCarthy test battery. In the memory test, the test administrator reads increasingly long sequences of numbers to the child, and asks the child to repeat them. The leg motor test measures the ability of children to execute six predetermined tasks—for example, walking on tiptoes or backwards, and standing on one foot.[12] In 2008 an associative memory test of the Woodcock-Muñoz test battery was also used.[13] In this test, children are introduced to an increasing number of space creatures with nonsensical names, and are asked to identify the different figures on a page that shows many of them at the same time. The final test we use is the Behavior Problem Index (BPI), which is based on the caregiver's report of the frequency that a child displays each of 29 problematic behaviors, with responses coded as "never", "sometimes" and "often".[14] We use the number of behavioral problems for which a caregiver answers "often". Unlike the other outcomes we study, behavioral problems do not necessarily indicate a delay, as there are no benchmarks or established ages at which they are predicted to decrease.[15]

Our analysis focuses on children below 6 years old when the transfers started. This means that all of them are below 7 years at the 2006 follow-up. The program requirement of primary school enrollment and attendance was binding for children age 7 and above. None of the children in our sample are bilingual—an obvious concern with tests that measure language ability. All of the tests were carefully pre-tested in the field and a handful of items that appeared to be culturally inappropriate were amended. The TVIP is standardized with a population of Mexican and Puerto

[11] Before the test starts, the enumerator explains the test with the help of a few example slides. She proceeds to the actual test slides only once the child has demonstrated understanding of the test.
[12] See Stoltzfus et al. (2001), Gertler and Fernald (2004), Fernald, Gertler, and Neufeld (2008), and Cogill et al. (1986) for other applications.
[13] The associative memory tests was added for comparability with other studies on ECD outcomes in Latin American countries, in particular Paxson and Schady (2010), Schady (2011) and Fernald, Gertler, and Neufeld (2008).
[14] Recent applications of the BPI in Latin America include Fernald, Gertler, and Neufeld (2008); Paxson and Schady (2010).
[15] There is some overlap between the BPI and the social-personal behaviors measured in the Denver. For instance, the Denver personal-social subtest has a number of items that relate to social interactions; and the BPI also has questions about whether or how the child interacts with others.

Rican children, and the words are all part of the vocabulary in Nicaragua. This was further verified during pre-testing.[16] In the case of the Denver test of language, there also does not appear to be an obvious concern with cultural appropriateness—the test measures whether infants can utter various sounds and, for older children, whether they can identify and name simple concepts, such as body parts. Additional evidence that the Denver is appropriate for our study population is provided by the fact that the national early childhood stimulation program in Nicaragua uses a slightly modified version of this test for child monitoring. For all these reasons, it seems unlikely that the observed delays are due to possible cultural inappropriateness of the tests for children we study.

All tests were administered by a specially trained team of female-only test administrators. They were selected for their background (training as psychologists, social workers, or similar) and for their ability to establish good contact with small children. During the training, high emphasis was placed on gaining the confidence of the children before starting the test administration and on the standardized application of each of the tests. Moreover, data collection and test administration was organized in such a way that the test administrators would maintain a balance between the number of children visited in treatment and control communities, and visits to treatment and control were also balanced in time, to avoid any seasonal differences. Consistent with this approach, the results are robust to controls for the identity of the test administrator and for the month the tests was taken.

Measurement of intermediate inputs

[16] Paxson and Schady (2007) show that in rural Ecuador children whose mothers or fathers have completed secondary schooling have average scores that place them in the 50th percentile of the test, indicating they perform as well as the international reference population, even if many others in the same setting also have very large delays. In our sample, only a very small number of parents (5.3 percent of mothers, and 4.5 percent of fathers) have completed secondary school, preventing us from carrying out a similar calculation. Nevertheless, it is telling that the results of the Denver reported in Oberhelman et al. (1998) for a Nicaraguan population that has much higher education levels show much smaller delays.

As in most surveys, measures of input use are obtained by caregiver reports (for child food intake and stimulus indicators), or by reports by the household head or his spouse (for household food consumption and health indicators). Given the social marketing component of the intervention, this raises potential concern regarding reporting bias, as informants in the treatment could, in principle, be more likely to provide answers that they believe are socially desirable. The social marketing did not include messages on early childhood stimulation, making it less likely that we would observe reporting bias on those indicators. Also, while reporting error might be of concern for the measurement of other intermediate inputs during the intervention (in 2006), such concerns a priori seem to be much less for the results for 2008, when the intervention had concluded for approximately two years. Moreover, for at least one measure, the self-reported measures of growth check-ups, we can double-check the information using a vaccination-health use card filled in by health care providers when children visit the health center. Enumerators directly observed the presence of those cards. In 2006, the share of children with cards is higher in the treatment group, consistent with the fact that the health use cards are given to parents during growth check-ups. Among children with cards, however, there is no difference between treatment and control groups in the extent to which the information derived from the cards is consistent with answers provided by respondents in the interview. Hence, the finding that a higher share of children went to growth controls in the treatment is unlikely to be due to overreporting of desirable behavior by treatment parents. For 2008, children in the treatment are no longer more likely to be weighed and, consistently, they are also no longer significantly more likely to have vaccination-health use cards. In 2008, too, the level of concordance between parental reports and the cards is not different between treatment and control groups.

Table A1: Baseline characteristics of all eligible households compared to eligible households tracked back at follow-ups

	Mean all	Mean all reinterviewed	
		2006	2008
Household characteristics			
male household head	0.82	0.82	0.83
household size	5.28	5.30	5.33
# hh members 0-5 years old	0.65	0.65	0.65
# hh members 5-14 years old	1.53	1.54	1.55
# hh members 15-24 years old	1.07	1.08	1.08
# hh members 25-64 years old	1.77	1.78	1.79
# hh members more than 65 years old	0.24	0.24	0.24
Number of rooms in the house	1.62	1.63	1.63
Time to school (minutes)	0.27	0.27	0.27
Time to health center (minutes)	1.14	1.14	1.14
Time to municipal headquarters (minutes)	1.54	1.54	1.54
Owns toilet/latrine	0.76	0.76	0.76
Access to water	0.13	0.13	0.13
Access to electricity	0.40	0.40	0.40
Own land	0.66	0.66	0.66
Total Consumption per capita (cordobas)	5443	5424	5410
Food consumption per capita (cordobas)	3537	3527	3524
Number of households	4021	3969	3923

Note: Sample includes all eligible households in treatment and control

Table A2: Baseline characteristics of eligible children who did all age-relevant tests compared to all eligible children

	2006		2008	
	All	Did all tests	All	Did all tests
Child-specific characteristics				
All children				
male	0.50	0.50	0.50	0.50
age in months when transfers started	32.80	32.45	21.10	21.38
mother lived in household at baseline	0.95	0.95	0.96	0.96
# years education mother	3.98	4.01	4.08	4.08
# years education father	3.65	3.67	3.82	3.83
Children age 3-6 at baseline				
tvip (vocabulary recognition) test score	6.01	6.05	6.00	6.02
Children age 0-5 at baseline				
weight-for-age z-score	-1.02	-0.98	-1.01	-1.01
height-for-age z-score	-1.23	-1.19	-1.22	-1.22
weight-for-height z-score	-0.16	-0.14	-0.17	-0.17
birth weight	6.79	6.81	6.79	6.79
weighed in last 6 months	0.91	0.91	0.91	0.91
Received vitamins in last 6 months	0.69	0.69	0.69	0.70
Received deworming drugs in last 6 months	0.53	0.52	0.53	0.54
Household-level characteristics				
male household head	0.85	0.86	0.85	0.85
household size	6.09	6.10	6.19	6.20
# hh members 0-5 years old	1.16	1.17	1.08	1.09
# hh members 5-14 years old	1.72	1.72	1.76	1.76
# hh members 15-24 years old	1.18	1.18	1.27	1.27
# hh members 25-64 years old	1.87	1.86	1.90	1.91
# hh members more than 65 years old	0.15	0.15	0.16	0.15
Number of rooms in the house	1.59	1.58	1.60	1.61
Time to school (minutes)	0.27	0.28	0.27	0.27
Time to health center (minutes)	1.20	1.21	1.20	1.20
Time to municipal headquarters (minutes)	1.63	1.63	1.61	1.62
Owns toilet/latrine	0.72	0.72	0.73	0.73
Access to water	0.12	0.12	0.12	0.12
Access to electricity	0.37	0.37	0.37	0.37
Own land	0.63	0.64	0.64	0.65
Total Consumption per capita (cordobas)	4548	4547	4561	4567
Food consumption per capita (cordobas)	3095	3098	3102	3111
Proportion of food in total expenditures	0.69	0.69	0.69	0.69
Proportion of staples in all food exp.	0.59	0.58	0.59	0.59
Proportion of animal proteins in all food exp.	0.16	0.16	0.16	0.16
Proportion of fruit and vegetables in all food exp.	0.05	0.05	0.05	0.05
Number of children	3333	3135	4316	4153
Number of households	2221	2137	2604	2539

Note: Sample is based on all children in each year visited by the test administrator that were either younger than 6 years when transfers started or born to baseline household members since the baseline.