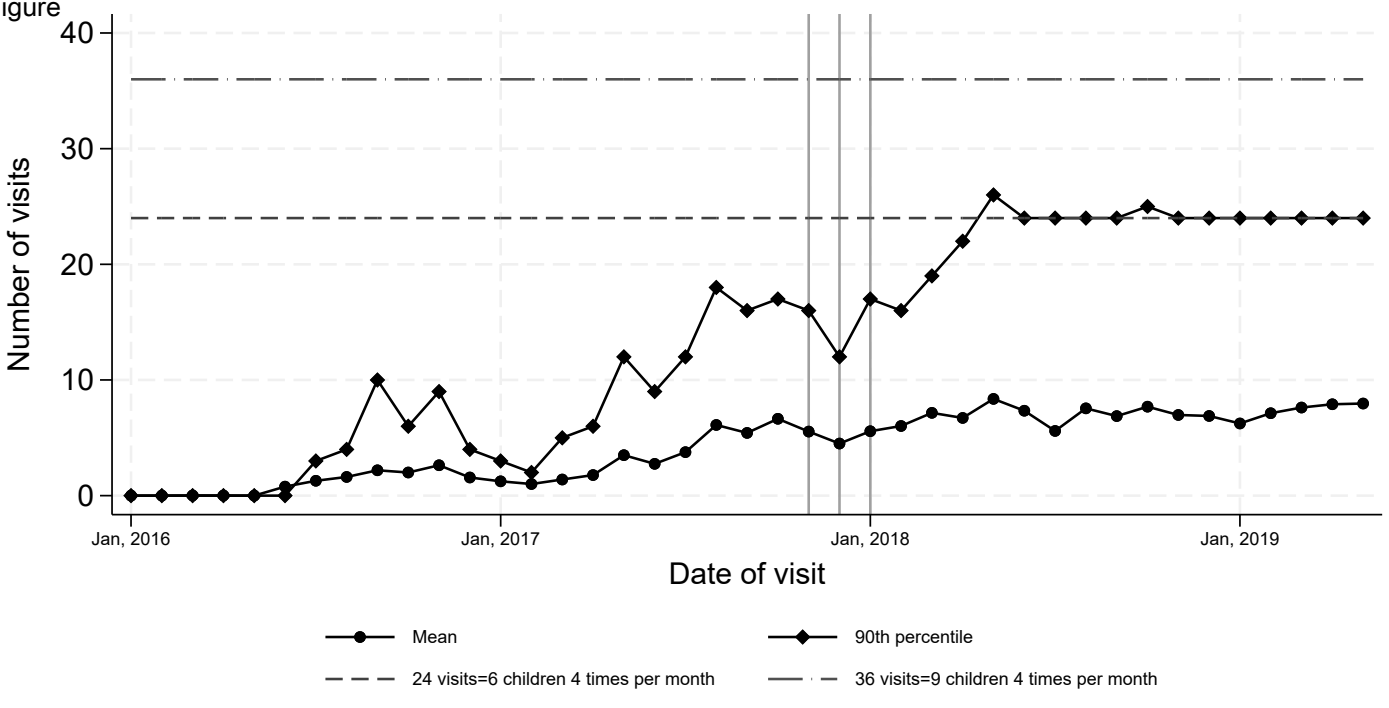
Journal of Human Capital

Impacts of Integrating Early Childhood with Health Services: Experimental Evidence from the Cresça Com Seu Filho Home Visiting Program

--Manuscript Draft--

Manuscript Number:	220143R3
Full Title:	Impacts of Integrating Early Childhood with Health Services: Experimental Evidence from the Cresça Com Seu Filho Home Visiting Program
Article Type:	Special Issue Submission - Childhood Human Capital Development
Keywords:	Estratégia Saúde da Família; community health workers; home visiting; early childhood development; Brazil
Corresponding Author:	Florencia Lopez Boo Inter-American Development Bank Washington, DC UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Inter-American Development Bank
Corresponding Author's Secondary Institution:	
First Author:	Florencia Lopez Boo
First Author Secondary Information:	
Order of Authors:	Florencia Lopez Boo
	Maria de la Paz Ferro
	Pedro Carneiro
Order of Authors Secondary Information:	
Abstract:	Delivering early childhood programs at scale is a major policy challenge. One way to do so is by using existing public infrastructure. This paper experimentally assesses the short-term impacts of a new government home visiting program integrated into health care services. The program changed the allocation of time for community health workers, asking them to carry out early childhood development-related tasks. We find that access to the program has a positive but modest impact on home environment quality and no impact on child development nor on children's health status. Our results point to the importance of workload, supervision and buy-in from delivery actors to enhance fidelity of interventions.

Figure



All CHW who conducted at least one full visit during the study period are included (N=172).
The vertical lines indicate the date the bonus system is implemented:
1. Nov 2017. Requirement: at least one eligible child surveyed by the ACS between Jan-2015 and Aug-2017.
2. Dec 2017. Requirement: at least one child visited between June-2017 and Sept-2017.
3. Jan 2018 to present. Requirement: start of performance-based pay plan. At least 6 to 9 children visited per month.

Impacts of Integrating Early Childhood with Health Services: Experimental Evidence from the *Cresça Com Seu Filho* Home Visiting Program[†]

Florencia Lopez Boo[a*], Maria de la Paz Ferro[a], Pedro Carneiro[b]

Abstract

Delivering early childhood programs at scale is a major policy challenge. One way to do so is by using existing public infrastructure. This paper experimentally assesses the short-term impacts of a new government home visiting program integrated into health care services. The program changed the allocation of time for community health workers, asking them to carry out early childhood development-related tasks. We find that access to the program has a positive but modest impact on home environment quality and no impact on child development nor on children's health status. Our results point to the importance of workload, supervision and buy-in from delivery actors to enhance fidelity of interventions.

[†] This research was made possible by the unconditional support of the Prefecture of Fortaleza and the Fortaleza Municipal Health Secretariat. We wish to extend a special thanks to Carolina Cunha Bezerra, Cristiana Ferreira da Silva, and the technical teams from the PCCSF and the ESF for their commitment to this experimental implementation and evaluation. The authors wish to thank Marcia Rocha (IADB/Brazil) for her tireless dedication to implementing and evaluating the program, and Jennie Encina and Mayaris Cubides for their research assistance. Pedro Carneiro gratefully acknowledges the support of the European Research Council through grant ERC-2015-CoG- 692349. We also would like to acknowledge the work of the team at the firm Overview during the data collection phase. Baseline and follow-up data collection was funded by two Technical Cooperations from the Inter-American Development Bank (BR T1300/BRT1330).

[a] Inter-American Development Bank

[b] UCL, IFS, CEMMAP, FAIR-NHH

* Corresponding author (email: florencial@iadb.org)

I. Introduction

There is broad consensus regarding the importance of early experiences and stimulation during early childhood (during the first five years of life) as determinants of better outcomes in subsequent stages of life (Heckman, 2008). Simple stimulation activities like playing, reading, and singing with a child can improve their ability to think, communicate, and connect with others (Black et al., 2017). Likewise, a warm and positive home environment is crucial if children are to reach their development potential (Caldwell, 1967).

To ensure children receive quality experiences during this critical stage of life, one public policy response has been to invest in early childhood (EC) development programs. Home visits by community workers to local families have been discussed as a particularly promising way of providing psychosocial stimulation, given their success in improving parenting practices and child development (Berlinski and Schady, 2015). In general, the aim of these visits is to encourage stimulating activities, good health practices, and positive disciplinary strategies within the home, resulting in an improvement in environment quality and, therefore, on children's cognitive and socio-emotional development.

In this paper we present results from an experimental evaluation of a home visiting program for vulnerable households with children ages 0 to 3 years old (*Programa Cresça com seu Filho*, or PCCSF) delivered through Brazil's primary healthcare system (the *Estratégia Saúde da Família*), and implemented at scale in the city of Fortaleza. The aim of PCCSF is to promote children's development through play-based activities, and the program differs from other early stimulation interventions in that it ensured the ability to scale up by being integrated into the *Estratégia Saúde da Família* (ESF), with extensive penetration guaranteed by the community health workers' (CHWs') 30 years of experience in the neighborhoods benefited by the program. Furthermore, the

curriculum is designed around this integration with primary care services, so the module aiming to improve parenting practices and child development was reviewed in conjunction with all other service priorities for the target population, such as promoting breastfeeding and immunization and preventing diarrhea in children.

In the evaluation sample, access to PCCSF was randomized across small geographic areas. We study experimental program impacts on child development and health, and on parental behaviors (cognitive stimulation; responsivity and behavior management; and investments in the child). We find that access to PCCSF (intent to treat, or ITT) had a moderate positive impact on home environment quality, of 0.12 standard deviations (SD), especially with regard to acceptance and behavior management, which are practices that a set of activities within the PCCSF curriculum sought to impact directly. However, we did not find the program to have had an impact on parents' activities that directly promote the child's learning and development (e.g., reading and playing with the child), nor did we find any evidence of substantial impacts on the child's development or health status.

It is possible that the lack of impacts on children is due to low quality in the program's quality implementation in terms of visit compliance, intensity, and fidelity to the intended visit design. In fact, families in locations with access to PCCSF experienced on average less than two visits more than families in locations without access to PCCSF, whereas in theory they should have received in excess of 40 visits more (over the course of a year) than those in non-PCCSF areas.

We estimate that the impact of one PCCSF visit on home environment quality (by instrumental variables, or IV) was 0.07 SD, which is quite large. Given the low number of visits in our sample it is difficult to extrapolate what would have happened if recipient families benefited from the scheduled number of visits, and it is likely that the marginal impact of visits declines with its

number. But given the very high estimated impact of one visit, it is reasonable to presume that the program would have had a large impact on home environment if it had been widely and frequently accessible to the target population.

Why was the actual number of visits so low in PCCSF areas? In the last sections of the paper, we present some hypotheses but the most plausible ones are: the workload of the workers was not taken into consideration when planning the intervention, there were not enough workers to manage the additional workload, and there were no adequate incentives for adopting a more intensive workload among the existing workers.

There is a large literature examining the impacts of home visit programs more generally (not only those integrated in national health systems in low- and middle-income countries). A meta-analysis of 40 programs in the United States suggests short-term effects on child development of 0.10 to 0.20 SD (Sama-Miller et al., 2017) and longer-term effects on cognitive achievement and behavior, for example, from the Nurse-Family Partnership (Olds et al., 2004, Olds et al., 1999), which is implemented by health workers. One of the most influential studies in a developing country is the now called Reach Up curriculum, first implemented in Jamaica by trained CHW. An originally small scale program, it showed large impacts on cognitive, socio-emotional, and health outcomes during childhood (Grantham-McGregor et al., 1991, 1994), academic performance, criminal behavior, and health during adolescence (Walker et al., 2011, Baker-Henningham et al., 2003), and earnings during adulthood (Gertler et al., 2014, Gertler et al 2021). Pilot home visiting programs like the one in Jamaica have had short-term impacts of 0.20–0.70 SD on child development and parenting practices (Andrew et al., 2020; Worku et al., 2018; Attanasio et al., 2014; Lozoff et al., 2010; Nahar et al., 2009; Hamadani et al., 2006). However, many

programs that show promise in tightly-controlled, small-scale settings fail to achieve similar effects at scale (List et al., 2021; Araujo et al., 2019; Lopez Boo et al., 2020).

However, because EC programs such as home visits are often so costly to implement, it has been recommended that they be integrated into existing health services (Black and Dewey, 2014; Chan, 2013),[1] which not only contributes to cost reductions, but also allows for a better exploration of complementary interventions fostering the health, cognitive and socio-emotional development of children (Hamadani, 2019, Richter et al., 2017). In the United States, for instance, programs implemented at pediatric offices with the aim to improve parenting practices had a modest but significant impact on encouraging positive parent-child interactions and cognitive stimulation activities (Shah et al., 2016).

There are far fewer home visiting programs implemented from within the health system with an emphasis on child development in vulnerable families in low- and middle-income countries. Experimental evaluations showed positive results on child development in Pakistan (Gowani et al., 2014; Yousafzai et al., 2014), Jamaica (Walker et al., 2015), China (Luo et al., 2019; Sylvia et al., 2020), and Gambia (Blimpo et al., 2016). The results were moderate in Pakistan, Gambia, and China. Additionally, the two studies from China and the one from Gambia found positive impacts on parenting practices. The literature suggests that the key factors for the effectiveness of these

[1] Traditionally, the primary goal of home visiting programs that are integrated into health systems has been to improve children's health, and these programs have had positive effects on this type of variable (Shakya et al., 2017; Cheng et al., 2019; Janmohamed et al., 2020). They are generally carried out via peer counseling or by health professionals, and their goal is to encourage early initiation of breastfeeding and exclusive breastfeeding (Ara et al., 2018), reduce infant mortality (Brener et al., 2011), promote growth (Nair et al., 2017), and increase variety in children's diets (Shi et al., 2010). Six months of visiting severely malnourished children who are admitted to a rehabilitation unit also shows positive impacts on the children's development (Nahar et al., 2009). In Brazil, a historic example of a program with significant impacts in terms of reducing child mortality, improving child nutrition, and expanding immunization coverage is the ESF. This program, which has had broad national reach since the mid-1990s, aims to continue expanding access to basic health services by sending health professionals to communities. These professionals identify risks, respond to emergencies, administer vaccines, and work to promote health within the home (NCI, 2019).

integration methods are workload, health workers' level of training, and supervision (Yousafzai, 2014; DiGirolamo et al., 2014; Walker et al., 2017). It is important to note that, unlike the PCCSF, none of the above-mentioned studies were implemented by government agencies nor at full scale. Implementation at large scale can hinder program's impacts because of the combination of low resources and the fact that a public program means that implementers usually do not want to leave any beneficiaries out, so there is a temptation to stretch resources to the point that they are ineffective (List et al, 2019).

In the related literature above, programs were either designed and/or implemented by research teams, in a limited number of locations and with few families.[2] The PCCSF, on the other hand, was implemented directly by the government with no supervision of the research team, in a big city in Brazil, reaching more than 40% of the eligible population in a more structured public health system. The evaluation of an at-scale, government implemented program is the central contribution of this paper.

This paper proceeds as follows. In section 2 we describe the intervention, the data, and the experimental design. In section 3 we show our main empirical results, and in the last section we conclude.

II. Intervention, data collected, and empirical strategy

A. Context and description of the Program

Brazil has been one of the most successful countries in Latin America and the Caribbean at reducing the mortality rate of children zero to five years old (Berlinski and Schady, 2015). This

[2] The exception is Blimpo et al (2016) in Gambia, which was implemented directly by the National Nutrition Agency, see Appendix C for a detailed description of these studies.

success is due in part to the universal health coverage Brazil achieved through a unified health system that offers community primary care (Bhalotra, Rocha and Soares, 2016; Hone et al., 2017): the ESF. The ESF, which is the gateway to the Universal Health System (Sistema Único de Saude—SUS) prioritizes health prevention and promotion, with the objective of improving the population's quality of life by monitoring people's health condition, detecting diseases at early stages, and educating the population to foster healthier habits. The ESF is composed of multidisciplinary teams that include: 1) a general or family health practitioner, 2) a general or family health nurse, 3) a nurse's assistant or technician, and 4) a CHW. CHW are assigned to geographical areas smaller than a neighborhood called a Micro Area (MA).[3]

The home visiting program PCCSF was launched in 2014 and now serves nearly 30,000 children under age 3 in Fortaleza, or about 40% of the total population in that age range.[4] It was developed through the ESF, with the backing of the Health Secretariat and Municipality of Fortaleza, and with the technical support of the Inter-American Development Bank, the Early Childhood Institute (Instituto da Primeira Infância—IPREDE), the Universidad Federal de Ceará (UFC), and the Better Early Childhood Program (*Primeira Infancia Melhor*—PIM). The PCCSF is a municipal initiative and started in three (out of seven) of the most vulnerable administrative divisions, to then be expanded to the rest of the city.

[3] CHW are government employees who work 40 hours per week. The number of CHW assigned to a given MA depends on population density. CHW working for the ESF usually do tasks related to health promotion, including guiding and supporting people who need care, conducting health related home visits, stimulating participation in health activities, and acting to control epidemics. CHW are required to have finished at least secondary school, take a mandatory course, and live in the MA where they will be working (at the time of the public request for applications).
[4]Fortaleza is the capital of the state of Ceará, in northeastern Brazil. It is the country's fifth most populous city (population as of April 2020: 2,659,138 (IBGE, n.d.)) and 10th richest city (in 2010, the poverty rate was 12.14%, which is the proportion of individuals with a per-capita household income of up to 140 Reais/26 USD per month). It is divided into seven regions, which are territorial divisions of the municipality (see map in Appendix A). The municipality of Fortaleza has devised several early childhood strategies, of which PCCSF is only one. Expanded enrollment in childcare, for example, has been one of the most successful strategies in recent years (see: https://forbes.com.br/brand-voice/2019/09/fortaleza-na-vanguarda-das-acoes-para-a-primeira-infancia/).

The PCCSF program draws its inspiration from the experience of the PIM (Verch, 2017; Ribeiro et al., 2018) and from the theoretical foundation laid out in *International Child Development Programmes*/*More Intelligent and Sensitive Child* (ICDP). It also draws on the Jamaican model *Reach Up and Learn* (RU), especially in the area of training (which was given by the RU team), supervising, monitoring, and mentoring CHW, and with regard to the "spirit" of the visit, that is, the relationship built between the visitor, caregiver, and child, and in how activities are demonstrated (Lopez Boo and Encina, 2015). Under the program's design, CHW trained in the PCCSF curriculum conduct weekly, one-hour visits with children ages 0 to 3 from vulnerable homes. Children are visited for two years or until they turn three.

Each visit has proposed activities for children in the age ranges: 0–6 months, 6 months to 1 year, 1–2 years, and 2–3 years. The activities were designed to promote the different child development domains (motor, language, cognition, and social-emotional) and increase in complexity as the child grows older. They particularly focus on best parenting practices, encouraging the caregiver to respond in a warm, sensitive, and responsive way to the child's signals, promoting positive discipline and stimulating the child's cognition and language. Additionally, at the beginning of the visit, the CHW and caregiver talk about the child's health (this includes requesting their Health Record and discussing good hygiene, nutrition, and accident prevention practices).

To monitor and reinforce the skills of CHW in their role as promoters of child development, a rigorous training system and a two-tiered supervision system (individual and group) were designed.[5,6]

Implementing the program According to the Program's administrative information, the PCCSF began being implemented in 2015 when CHW registered eligible children. Children are eligible if they are 0 to 28 months old,[7] the child or mother is registered in CadUnico registry[8], and they belong to a household with a monthly income of up to 500 Reais/120 USD. Registration of eligible children and the start of visits did not occur at the same time in all regions or within each region. In fact, when follow-up data was collected for this evaluation, there were still some CHW included in the evaluation that should have started conducting visits three years ago and that had not yet conducted a single PCCSF visit. In the first half of 2016, the first PCCSF visits were conducted in the evaluation region, and as implementation progressed, a trend of low CHW adherence to the program began to emerge. In response, the local government worked on incentive schemes, which are described in the next section.

[5] CHW and nurses in treatment MAs received parallel training on the curriculum and structure and quality of visits, as well as on the instruments and methods for targeting and monitoring, including the PCCSF supervision logs and a checklist for observing visit quality. The main aim while training CHW was to create opportunities during training to do practice visits and learn methods to deal with the different situations that could arise during visits. Meanwhile, training for nurses mainly focused on practical training on the supervision system.

[6] Group supervision consists of weekly meetings between the CHW and nurse to go over visits conducted the previous week. These meetings cover: difficulties when conducting visits, questions about using the handbook, and tracking the children's activities and progress. Individual supervision consists of a nurse going along on a home visit. The nurse is an observer and cannot participate in the visit. It is recommended that both levels of supervision be done weekly.

[7] This ensures the child will be visited for at least 8 months, the minimum time the literature finds as necessary in order to achieve an impact (see Baker-Henningham and Lopez Boo, 2010, among others).

[8] CadUnico is the abbreviation of *Cadastro Único para Programas Sociais do Governo Federal* (Centralized Registry for Social Programs of the Federal Government), a system that registers people and families in a state of social vulnerability due to low income. It is used to check whether people meet the requirements for social programs like the *Bolsa Família* transfers program.

According to the Program's calculations, the monthly per-beneficiary cost of the PCCSF is 83.67 Reais, or 14.71 USD (176 USD per year). This amount is based on costs from human resources (pay for CHW, nurses, data entry worker, and others), infrastructure, and materials. This cost is slightly higher than that of similar programs like the PIM (124 USD per year according to Versch, 2017) or the *Amor para los más chiquitos* visiting program (100 USD per year, Lopez Boo, Leer and Kamei, 2020), but it is substantially lower than programs like the pilot home visiting program *Familias en Acción* in Colombia (USD 500, Attanasio et al., 2014).

B. Study design

The experimental evaluation of the PCCSF was carried out in Region V (population as of April 2020: 591,903) of the municipality of Fortaleza, where 480 small areas (MA) were randomly assigned—half to treatment and half to control (see Figure 1, which summarizes the design of the study). There are an average of 1.04 CHW per control MA and 1.06 CHW per treatment MA. Each CHW is assigned to a single MA, so randomization at this level results in *de facto* randomization at the CHW level. Treatment CHW were expected to devote 16 hours of their 40-hour work week[9] on PCCSF-related activities, which include visiting a minimum of nine children per week, recording the visits in the system, and participating in the supervision session with the nurse. CHW in control MA continued on with their usual community-level health promotion tasks for the ESF. Therefore, the comparison in this paper is between two different allocations of the same amount of working time between CHW workers, and the question is whether it is productive to shift time worked by CHWs away from health promotion and towards home visits. It is possible that such a

[9] In 2018, the workload of CHW (both treatment and control) dropped to 32 hours to match the workload of other ESF health professionals.

shift produces positive or negative impacts on children and parents, depending on which type of activity is more productive.

Through this randomization, our research design allows us to estimate the Program's causal effects. Additionally, the ability to randomize at the level of smaller geographical areas would give us greater homogeneity within clusters (in this case, the MA) compared to previous studies that use broader geographical areas such as communities (Lopez Boo, Leer, and Kamei (2020) or Attanasio et al. (2014)).

Study design challenges: analysis of administrative data

One challenge for the evaluation was low CHW adherence to the program. It is possible that CHW were not willing to take on a change in the composition of their activities, and carried on doing the health promotion tasks they had been traditionally engaged in. Based on administrative data, 31% of CHW in treatment MA had not conducted a single Program visit as of June 2019. Even within the group of CHW who during the study period completed at least one full visit, implementation was low in terms of number of children visited and number of visits. Figure 2 shows the average number of visits conducted by a CHW for each month (circle-marker line), as well as the 90th percentile of number of visits for each month (diamond-marker line). While the number of visits rose over time, the average remained lower than expected in the program's original design (36 visits, or 9 children visited per week), and below the adjusted range following changes to the design two years after the program began (24 visits, or 6 children visited per week).[10]

[10] We see a similar pattern in the number of children a CHW visited each month: the number rises over time but is lower than the initial target (9 children), and below the adjusted range established in 2018 (6 to 9 children).

Given the low CHW adherence, the number of children to be visited was relaxed from the original minimum of 9 to a range of 6 to 9. Also, the program purchased tablets and created a system to distribute them as a reward to the best CHW; created a full-time position to support implementation; and, together with the Health Secretariat, devised a system of performance-based bonuses for CHW. This bonus system was based on CHW performance in terms of how many children they visited in a month and how many times they visited them.

The lower horizontal dashed line in Figure 2 represents the cutoff for receiving a bonus. The dates after which the bonuses were introduced are marked in Figure 2 by vertical lines. It is unclear whether the bonus system actually motivated the average CHW in any significant way. However, it does seem that the system managed to motivate "borderline" CHW (that is, those conducting an above-average number of visits who were close to meeting the requirements for receiving the bonus) to hit the minimum requirement for receiving the bonus.

For the 61 CHW in treatment MA -out of 252- for which we have information, we explored what characteristics of the CHW explain the number of visits carried. We check for CHW's gender, age, education level and experience as CHW, and if she participated in the PCCSF Program's training, if she completed the training, and if she believes the Program benefits the mother, child or CHW (see Table B1). We find that participating and completing the Program's training and having more than 10 years of experience are consistently correlated with the number of visits completed.

C. Questionnaires and data

In the second half of 2015, all children under age 3 who at the time were registered in CadUnico, were living in Region V, and lived in a household with a per-capita income of less than 500 Reais were contacted by the survey firm. The firm confirmed basic characteristics of the household and

that the child did indeed live in Region V. Based on power calculations, it was determined that a minimum of seven children should be surveyed per MA.[11] During the first semester of 2016, the children to be interviewed were randomly selected from this "census".[12]

Baseline data (before the CHW were to begin PCCSF visits on June 3, 2016) was collected between February and July 2016. The final baseline included surveys from 2292 children living in one of the evaluation's MA.[13] Follow-up information was collected 3 years later, between March and July 2019. A total of 1961 children (of the 2292) were evaluated during follow-up, for an attrition rate of 14% (15% in treatment MA and 14% in control MA). This attrition rate is a weakness of our data. Nonetheless, we do not find attrition to be associated with assignment to the treatment group (see Table B2) nor does differential attrition seem to be different between treatment and control (see Table B3).[14] To address concerns with attrition, we estimate Lee bounds (Lee, 2009).

For both the baseline and follow-up, we collected information on the child (including his or her health status and childcare attendance) and home, characteristics of the household members and child primary caregiver, and a rich set of measures of parental investment. We also administered a subset of items of the Early Childhood HOME Inventory (EC HOME) (Caldwell and Bradley,

[11] Prior to obtaining the baseline, assuming an intra-cluster correlation of 0.2, 7 children per MA, and 360 MA, we calculated the minimum detectable effect of the study without taking into account the possible efficiency gains of including controls in the estimation. With 80% power and for a two-tailed hypothesis test of size (α) equal to 0.05, we can detect an effect of 0.166 standard deviations.

[12] Apart from the three criteria the Program uses to identify potential beneficiaries, a fourth criteria was added to select the sample to be used in the study: children with any type of visible or diagnosed disability were excluded, since the literature suggests that these children need a different type of more focused intervention to achieve sustained improvement in their development (McCormick et al., 2006).

[13] 463 children who, while part of the baseline, live in a MA not part of the evaluation are excluded.

[14] We find there is differential attrition for 4 of the 25 characteristics: mother works, mother consumed drugs or alcohol during pregnancy, if the father lives in the home, and household size.

2003), and the Portuguese version of the Denver II Development Screening Test (referred to as the Denver throughout this study) (Frankenburg et al., 1990). The HOME measures the home environment quality, focusing on process variables and it is scored by observing interactions (not by caregiver self-reporting). The Denver assesses different areas of development (personal-social, fine motor, gross motor, and language) of children from birth until age 6.

For the analysis, we sorted variables into two groups; intermediate variables, which do not directly measure the child's development but are related to it, and final variables, which directly assess the child's development. We ended up with two final variables—a child development factor and an indicator for good child health—and three intermediate variables—two factors that assess home environment quality and one that measures investments in the child. To construct the development factor, we performed an exploratory factor analysis using continuous scoring internally normalized for age from the subscales for personal-social, fine motor, and language from the Denver. This analysis yielded a single factor with an eigenvalue greater than 1 and a factor loading for each subscale greater than 0.4, which we named the child development factor. Lastly, we standardized it with respect to the control group. We constructed the indicator for good child health using three questions about the child's health status that the caregiver had to answer (whether the child had any health problems, illness, or accident in the past 30 days; whether the child had diarrhea in the past 30 days; and whether the child had any illness accompanied by a high fever in the past 30 days). The child is considered to be in good health if the answer to these three questions was "no."

We grouped the survey variables assessing home environment quality—that is, the variables from the HOME module and disciplinary strategies modules—into one block, and variables that measure time investments parents made in their child in another block. For the variables measuring environment quality, we performed an exploratory factor analysis. This analysis showed that in

our data, the subscales learning materials, language stimulation, learning stimulation, and variety of the HOME measure a common underlying construct. These subscales evaluate the availability and variety of toys and activities to which the child is exposed that directly promote his or her learning and language development, as well as parent involvement in these activities. The other HOME subscales—responsivity, acceptance, and modeling—, together with questions about negative disciplinary strategies, evaluate a different construct associated with responsivity, sensitivity, and managing the child's behavior.[15] Based on this grouping, we constructed two factors. We termed the first group Cognitive Stimulation and the second group Responsivity and Behavioral Strategies. Using the information gathered to evaluate parent's investments in their children—time primary caregiver spends caring for the child, time the child spends watching television, simple activities like reading or telling stories that an adult does in the home with the child—we constructed a factor we termed Investments in the Child. Lastly, we standardized it with respect to the control group.

Appendix D provides detailed information on the questionnaires and instruments, and on the different independent variables used in the analyses, how they were constructed, and gradients due to socioeconomic characteristics of households.

D. Empirical strategy

To find the intention-to-treat effect, we estimated the following equation:

$$Y_i^t = \beta_0 + \beta_1 ITT_i + \beta_2 Y_i^{t-1} + \beta_3 X_i^{t-1} + \varepsilon_i \quad \textbf{(1)}$$

[15] We did not include the physical environment subscale of the HOME in either of the two factors because it did not correlate well to either. This subscale measures physical characteristics of the space.

Where Y_i^t is the outcome variable for child i at time t (follow-up); ITT_i is an indicator equal to 1 if child i lived in an MA assigned to the treatment group at time $t-1$ (baseline) and equal to zero if child i lived in an MA assigned to the control group at time $t-1$; Y_i^{t-1} is the same outcome variable for the child measured at $t-1$ (or a variable from the same development domain if Y was not measured at $t-1$; X_i^{t-1} is a vector of basic characteristics of the child and any imbalanced socio-demographic variable in baseline. Standard errors are clustered at the MA level. Given the random assignment of MA to treatment and control groups, the coefficient β_1, estimated using ordinary least squares, expresses the causal effect of making the Program available in an area relative to the *status quo* (intention to treat).

Since not all children surveyed in treatment MA received PCCSF visits, we also estimated the Program's effective impact. More specifically, only 11% of children surveyed during the follow-up period who were living in a treatment MA received at least one PCCSF visit. In the control MA, no children received PCCSF visits, meaning treatment compliance was perfect in this group. Thus, to find the Program's effect on the treated, we estimated a version equivalent to equation (1) using an instrumental variable method. The variable ITT was replaced with a continuous variable equal to the number of PCCSF visits the household effectively received. We obtained the information on visits directly from the Program's administrative visits records system. We used the random assignment to treatment (ITT) as an instrument for participation, and (because there is no contamination in the control group) the resulting estimate is interpreted as treatment on the treated (TOT).[16]

III. Results

[16] We evaluated the determinants of effective participation in the PCCSF including random assignment to treatment and control and socio-demographic indicators (see Table B4).

Implementation

Table 1 shows the distribution of the number of visits to children in treatment MA. Only 11% of children in treatment MA received at least one visit, and, within that group, 31% received more than 20 visits (the average was 15.83). In interventions similar to the PCCSF (in terms of mode, age of beneficiaries, emphasis of the intervention, and characteristics of the personnel conducting the visits) that found significant effects on child development, children received between 16 and 24 visits (Sylvia et al., 2020; Luo et al., 2019; Walker et al., 2015; Youzafzi et al., 2014; Gowani et al., 2014). In these studies, not only were the effects measured immediately after visiting ended (while in our case the average amount of time between the last visit and the follow-up measurement was on average 20 months), but also they all assess programs that are not at scale, which could lead to smaller effects in our case.

Additionally, the number of visits for children who did receive them was smaller than the number of visits envisioned in the program's original design. On average, depending on the child's age, they received between 36% and 60% of the expected number of visits. In short, our sample is characterized by a low number of children who received a visit and, for those who did receive at least one visit, a relatively low intensity. This surely dilutes any positive or negative impact that the program may have.

Baseline characteristics of the sample

Table 2 presents the baseline characteristics of the children and their homes for the sample evaluated in the follow-up period, by treatment status. The table includes basic characteristics of children and their mothers, variables related to time and money invested in the child, characteristics of the households, and outcome variables for children. We only found statistically significant

differences in two variables: whether there was an episode of armed violence in the neighborhood during the last month as reported by the household (51% in control MA, 45% in treatment MA) and whether the household received ESF visits (48% in control MA, 55% in treatment MA). This result suggests that the two groups of children are very similar and that the sample is balanced.

Effects of the program

Table 3 presents the effects of the PCCSF. Column one contains the intention-to-treat effects and column three the effect of effective treatment. We found a significant intention-to-treat effect on the factor of responsivity and behavioral strategies (an increase of 0.12 SD). The effect remains statistically significant after adjusting for multiple hypothesis testing using the procedure in Romano and Wolf (2005). We did not find the program to have had an effect on the other intermediate variables or on the measurement of the child's health or development.[17] In the last section of this paper, we discuss these results and possible reasons for the lack of impact on children. To put these results into context with the existing literature, the magnitude of our impact on parenting practices of responsivity and discipline is practically half of the effects found in both Luo (2019) and Sylvia et al. (2020), both programs designed and implemented under the direct supervision of the researchers and therefore not at scale. It is very much in line, however, with the modest impacts found in Blimpo et al (2016), which has however a program implemented in very particular conditions (by volunteers in a very rural and poor setting in Sub-Saharan Africa).

[17] The effects for each of the Denver subscales are also small and statistically non-significant (see Table B7). Furthermore, given the Program's curriculum emphasis on promoting child development through stimulating play-based activities, we ran additional analysis on different measures of play activities but did not find any significant effects (see Table B5).

As a robustness exercise, and to improve the precision of the coefficients, we estimated equation (1) using ordinary least squares: (i) without controlling for socio-demographic characteristics, (ii) controlling for the same set of variables excluding the baseline child development measurement, and (iii) the same controls plus fixed effects of the Primary Healthcare Unit (UAPS), given that the literature indicates that differing leadership style (in this case within the UAPS) would lead to very different service management styles. Regardless of whether controls were included and of which controls were included, the results are similar in both magnitude and significance, which can be explained by the baseline's excellent balance. We also estimated Lee Bounds for the effect of intention to treat on outcomes for which we had found a statistically significant effect. The bounds are different from zero (see Table B6).

Effective participation in the program impacted the responsivity and discipline strategies, as was the case for the ITT. In this case, we found that receiving one more visit has an impact of 0.07 SD on the responsivity and behavioral strategies factor, an impact of substantive magnitude. This suggests that the low ITT impacts are likely due to low program exposure, not to low program impacts.[18]

We also estimated other instrumental variables models where we changed the definition of participation in the program. More specifically, we estimated the effect of having received at least one PCCSF visit and receiving more than 20 full visits. The effects are robust to the definition of effective participation in the Program, and their significance remains the same in all cases.

[18] We estimated the effects on the individual components of the factor to check the effect was not being driven by the self-reported indicator (see Table B8). We find statistically significant effects on the Home Acceptance and Home Modeling subscales (although only the former remains statistically significant when using the Romano and Wolf correction).

We reviewed whether there were heterogeneous effects for characteristics of the children and families, including children's gender, age, and language development; the mother's level of education; home environment quality; and the caregiver's number of symptoms of depression. We did not find consistent heterogeneous effects for any of these characteristics.[19]

IV. Discussion and conclusions

As far as we know, there is little empirical evidence about the effectiveness of programs that integrate an early childhood program using psychosocial stimulation with health promotion programs at scale. In this paper, we evaluate the short-term effects of the PCCSF, a home visiting program implemented through the ESF in Fortaleza that currently serves 30,000 children under age 3. Since CHW participation in the program gradually increased over time, the effects we evaluate are thus interpreted as the average effects since the start of the Program for a recently launched program.

Our results show that access to PCCSF had a moderate positive impact on home environment quality, particularly on caregivers' sensitivity, responsivity, and behavioral strategies. This resulted in, for example, better use of positive disciplinary strategies instead of negative ones, increased parent ability to accept and handle negative behavior from the child, and enhanced parent ability to communicate expectations. A set of activities in the PCCSF curriculum were designed with the primary objective of encouraging good parenting practices, setting limits in a positive way,

[19] Prior to collecting follow-up data, the administrative information on PCCSF visiting indicated that implementation had been low and heterogeneous, so it was decided to mitigate the risk of low compliance in our sample by collecting an additional cross-sectional sample with potentially greater statistical power that would allow us to perform a parallel, non-experimental evaluation. We thus collected a matched control sample from treatment MA. A discussion of this sample and how it was used is available upon request. We're not presenting the results in this paper because of their methodological limitations.

explaining rules rather than just prohibiting, and making expectations clear. In other words, the specific aim was to directly impact the variables where an effect was indeed found.

Given the curriculum's emphasis on promoting child development through stimulating, play-based activities, the small and statistically insignificant effects on parental involvement in encouraging activities that directly promote children's learning and development (measured by the cognitive stimulation factor and the investments in the child factor) are surprising.

We also found small and statistically insignificant effects on child development.[20] Evidence from home visiting programs suggests that impacts of this type of interventions on children come about through changes in the home, not through direct stimulation during visits (J-PAL, 2020). While the Program had positive impacts on parental disciplinary practices, we did not find any effect on investments of time or resources in play activities, which could explain these results on child development. Our estimate aligns with a meta-analysis that finds that even if home visits increase parental knowledge and investments in children, they do not always necessarily impact child development (Supplee and Duggan, 2019). It is also in line with the meta-analysis by Shah (2016) that finds interventions based on pediatric consultations in the United States to have had moderate impacts on parenting practices. We also found that the magnitude of our results is in line with a similar intervention implemented at scale in Gambia by volunteers in an extremely poor setting (Blimpo et al., 2016) and is less than half of the impact of similar interventions implemented on a small scale and with the research teams rather than government agencies in charge of implementation.

[20] We have to acknowledge that our paper does not address optimal investment by parents and how that could be affected by the program's inputs like in Ehrlich and Yong (2013), something that should be explored in future research.

These results are also what could be expected from an intervention with poor compliance with treatment assignment, where the visits do not change investments of time and resources in children but do change some parenting practices, although not enough to impact child development. CHW focus groups reported that violence in MA (which made it difficult to conduct visits in certain areas), together with outbreaks of viruses like measles, dengue, and chikungunya, increased the workload of CHW, so they had to skip some PCCSF visits (IDB, 2019).

The low intensity of treatment offers a potential explanation for this result: 69% of children who received at least one visit were visited less than 20 times, and almost 3 weeks would pass between visits (mean = 2.8; standard deviation = 1.9). This runs counter to the original design, which established weekly visits in line with what the literature shows to be effective (Powell and Grantham-McGregor, 1989). Given that we estimate that the impact of one visit on home environment quality is 0.07 SD, it is likely that the impact of 40 or more visits (over the course of a year) on this outcome would be substantial, even if not a full 2.8 SD (0.07 SD * 40). These could then materialize into important improvements in child developmental outcomes.

Another factor for low impacts could have been low fidelity: visits lasted an average of 39 minutes instead of the full hour established in the design.[21] Additionally, in a focus group study we conducted with different PCCSF stakeholders (IDB, 2019), participants mentioned that parents often wanted to talk about other matters, so it is also possible that a percentage of the visit was spent on topics not strictly related to the child, reducing the likelihood of impacting the desired outcomes. We also know it was difficult to satisfactorily hold program visits because when CHW would arrive at the home, the family would not be available (whether because they were not there

[21] However, it is important to note that the PCCSF did achieve the same visit duration as visits in Jamaica, which did manage to change ECD outcomes (Walker et al., 2015).

or because the child was asleep). Both low fidelity and low intensity of implementation are two of the main challenges in scaling up social programs (List et al, 2021).

Finally, it should be noted that our study has the statistical power to detect plausible effect sizes given the literature on short-term impacts at scale (Araujo et al., 2019; Lopez Boo et al., 2020). In fact, the original power calculations postulated a short-term impact of 0.15 SD, while the analysis with the follow-up sample showed that an impact of 0.168 SD could be detected with 80% power.[22] While we cannot reject the possibility that the lack of impact could be due to the lack of statistical power, the effects on the child development factor are very close to zero. This magnitude is so small that our interpretation is that treatment made no difference in a child's development trajectory.

Beyond the results of the evaluation, this paper provides some lessons for scaling up social programs; and in doing so, we address an urgent gap in the early childhood policy literature regarding effective implementation of interventions at scale (List et al., 2021). First, while implementing home visits as part of the health system can allow EC interventions to be scaled up so they can reach a greater number of children, this evaluation reveals the challenge of using professionals who are already doing other tasks and who have no specific training in early childhood to conduct visits that aim to stimulate child development. In fact, in our analysis of focus groups with different PCCSF stakeholders, some CHW mentioned their excessive workload and the difficulty of adding the Program visits into their work routine (IDB, 2019). Second, it is

[22] For example, Lopez Boo et al. (2020) find a positive impact of ~ 0.15 SD on child development two years after the end of a home visiting intervention in Nicaragua, while Araujo et al. (2019) find an impact of 0.10 SD, also two years after the end of the Cuna Más home visiting in Peru). In our case, for the size of the follow-up sample size and an intra-cluster correlation of 0.2 (assumption used in the initial calculation), with 80% power and for a two-tailed hypothesis test of size (α) equal to 0.05, we can detect effects of 0.168 SD.

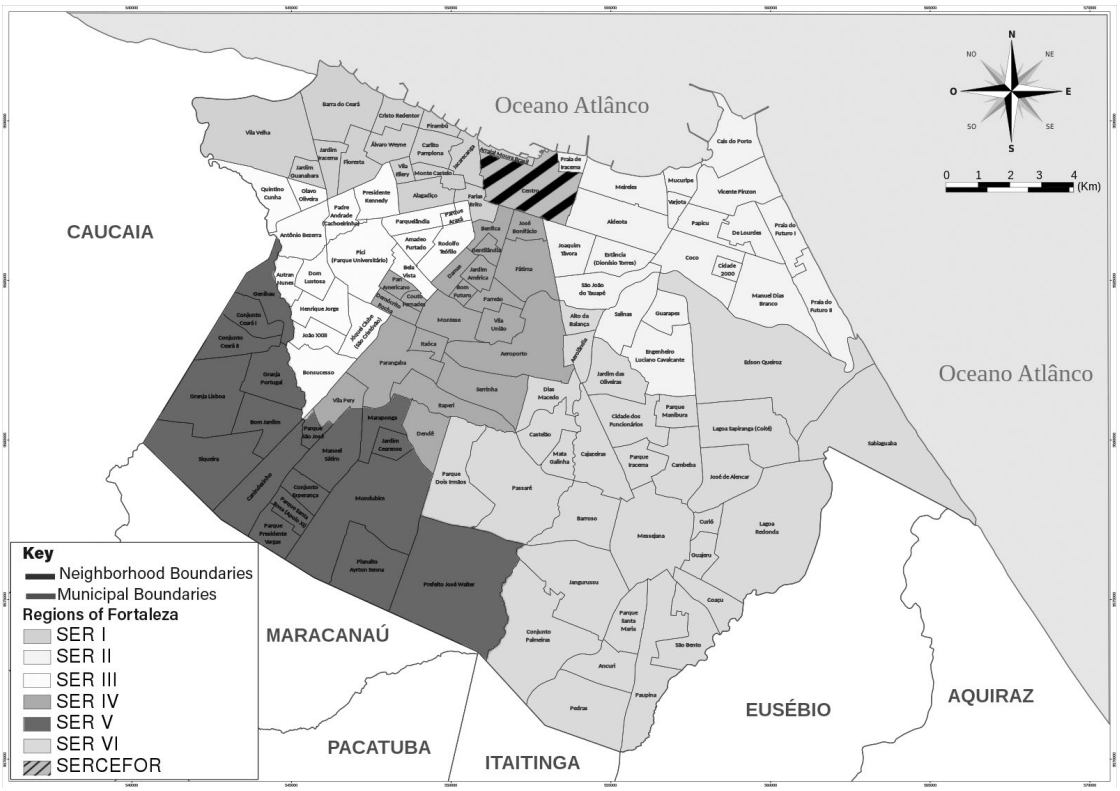
important to mention that while introducing the bonus system did lead to improvement in the Program's implementation, the change is not substantial. The post-bonus increase in number of visits mainly occurred in CHW who were already conducting an above-average number of visits but had not yet hit their target (see Figure 2). It is also important to mention that we did not find the Program to affect the likelihood of a household with children 0-3, being visited by a ESF's CHW (see Table B9). This result suggests that the Program did not displace the preexisting service within the PCCSF target population, which is important when considering its benefit.[23]

[23] One limitation of our data is that we do not have information to measure spillovers in the MA.

Appendixes

A. Map of Fortaleza

Figure A1. Map of Fortaleza



(Source: Wikimedia Commons. (2017). *Fortaleza Neighborhoods and Divisions* [Map]. Retrieved November 15, 2021, from https://commons.wikimedia.org/wiki/File:Bairros_e_divis%C3%B5es_de_Fortaleza_%28frame%29.svg)

B. Additional tables

Table B1. Correlations between CHW characteristics and PCCSF visits

	N of visits		At least 1 visit		>P50 visits		>P75 visits	
	Coeff	p-value	Coeff	p-value	Coeff	p-value	Coeff	p-value
Gender (male=1)	-23.66	0.62	0.01	0.93	0.01	0.96	-0.05	0.70
Age	1.54	0.64	0.01	0.46	0.00	0.84	0.00	0.86
Education level (completed secondary school or more=1)	17.50	0.82	0.07	0.69	0.13	0.48	0.14	0.30
Education level (more than secondary school=1)	44.09	0.33	-0.01	0.93	0.23	**0.09**	0.18	0.14
Experience (years)	4.51	0.35	0.00	0.84	-0.00	0.95	0.01	0.17
Experience (10 years or more)	88.94	**0.02**	0.24	**0.08**	0.18	0.20	0.20	**0.05**
Participated in PCCSF training	129.79	**0.00**	0.42	**0.02**	0.45	**0.00**	0.29	**0.00**
Completed PCCSF training	100.68	**0.04**	0.34	**0.02**	0.50	**0.00**	0.25	**0.01**
PCCSF has benefits (>0) †	17.10	0.72	0.10	0.41	0.12	0.35	-0.08	0.52
PCCSF has benefits (Nb) †	5.15	0.71	0.04	0.15	0.03	0.53	-0.01	0.71

Each cell corresponds to a separate OLS regression with robust standard errors. The dependent variable is in the first row. P-value<0.10 in bold. N= 61 CHW in treatment MA. †CHWs were asked if they believed the PCCSF had any benefits for the child, the mother, the family, or the CHW. Their answers were codified into categories. "PCCSF has benefits (>0)" is an indicator of whether the CHW states the PCCSF has any benefits, and "PCCSF has benefits (Nb)" is the number of benefits (categories) the CHW mentions.

Table B2. Evaluating determinants of attrition

Baseline socio-demographic indicators	N	Surveyed at follow-up		Not surveyed at follow-up		P-value
		Mean	SD	Mean	SD	
Assignment to treatment or control group	2,292	0.496	0.500	0.505	0.501	0.776
Characteristics of child						
Age (in months)	2,292	18.828	5.840	18.477	5.789	0.277
Gender (male=1)	2,292	0.515	0.500	0.465	0.500	**0.093**
Attends a childcare center	2,292	0.094	0.292	0.112	0.316	0.364
Characteristics of mother						
Years of education	1,884	9.469	2.845	9.278	2.832	0.292
Works	2,186	0.380	0.485	0.301	0.459	**0.004**
Number of symptoms of depression†	2,186	8.666	7.625	9.750	8.100	**0.028**
Consumed alcohol/drugs during pregnancy	2,287	0.067	0.250	0.094	0.292	**0.092**
Investment in child						
Investments in child (factor)	2,292	-0.021	0.962	0.031	1.097	0.422
Children's books at home	2,265	1.254	3.275	0.976	2.500	**0.069**
HOME: Responsivity	2,292	1.365	1.447	1.384	1.580	0.832
HOME: Acceptance	2,292	4.561	0.918	4.456	0.991	**0.063**
KIDI ‡	2,292	8.618	1.502	8.668	1.656	0.620
Characteristics of the household						
Father lives in the home	2,292	0.680	0.467	0.680	0.467	1.000
Household size	2,292	4.307	1.507	4.272	1.560	0.709
Per-capita income (R$)	2,236	257.273	206.367	241.270	194.016	0.167
Bolsa Familia Beneficiary	2,292	0.789	0.408	0.792	0.407	0.915
Armed violence in neighborhood (last month)	2,292	0.477	0.500	0.508	0.501	0.310
Violence in neighborhood	2,292	0.745	0.436	0.722	0.449	0.403
Household is visited by ESF	2,292	0.513	0.500	0.453	0.499	**0.054**
High perception of health service	2,292	0.505	0.500	0.495	0.501	0.763
Child outcomes						
Denver: Personal-social	2,292	14.690	4.030	14.423	4.024	0.235
Denver: Fine motor	2,292	16.881	3.295	16.825	3.296	0.750
Denver: Language	2,290	17.915	4.823	17.205	4.905	**0.011**
Denver: Gross motor	2,289	20.103	3.935	19.734	4.180	0.142
Good health	2,292	0.365	0.482	0.347	0.477	0.542

OLS estimation correcting standard errors by MA cluster. The dependent variable is equal to: 1 if the child could not be surveyed during follow-up and 0 in the opposite case. P-value<0.10 in bold. †Measured using the CES-D self report scale, 12 items (Radloff, 1977). If a person's score is equal to or greater than 10, they are considered to have significant symptoms of depression. ‡A reduced 12-item version of the Knowledge of Infant Development Inventory (MacPhee, 1981) was administered. This measure was only collected at BL.

Table B3. Testing for differential attrition

	N	Treatment assignment		Characteristic		Treatment assignment * Characteristic	
		Coeff	P-value	Coeff	P-value	Coeff	P-value
Characteristics of child							
Age (in months)	2,292	-0.013	0.783	-0.002	0.266	0.001	0.691
Gender (male=1)	2,292	-0.007	0.736	-0.035	0.099	0.022	0.441
Attends a childcare center	2,292	0.013	0.442	0.063	0.101	-0.082	0.123
Characteristics of mother							
Years of education	1,884	0.036	0.507	-0.001	0.725	-0.003	0.554
Works	2,186	0.025	0.228	-0.018	0.402	-0.049	**0.099**
Number of symptoms of depression†	2,186	-0.009	0.699	0.001	0.334	0.002	0.377
Consumed alcohol/drugs during pregnancy	2,287	-0.009	0.569	-0.034	0.322	0.169	**0.003**
Investment in child							
Investments in the child (factor)	2,292	0.004	0.781	0.017	0.176	-0.021	0.229
Children's books at home	2,265	0.002	0.885	-0.005	0.024	0.003	0.434
HOME: Responsivity	2,292	0.010	0.630	0.003	0.706	-0.004	0.698
HOME: Acceptance	2,292	-0.024	0.758	-0.018	0.102	0.006	0.708
KIDI ‡	2,292	-0.091	0.324	-0.003	0.691	0.011	0.295
Characteristics of household							
Father lives in the home	2,292	-0.037	0.167	-0.031	0.217	0.061	**0.071**
Household size	2,292	0.091	0.052	0.007	0.318	-0.020	**0.040**
Per-capita income (R$)	2,236	-0.007	0.786	-0.000	0.107	0.000	0.462
Bolsa Familia Beneficiary	2,292	0.021	0.526	0.013	0.611	-0.021	0.564
Armed violence in neighborhood (last month)	2,292	0.002	0.932	0.012	0.588	0.007	0.807
Violence in neighborhood	2,292	-0.006	0.844	-0.021	0.423	0.013	0.706
Household is visited by the ESF	2,292	-0.005	0.844	-0.041	0.049	0.022	0.473
High perception of health service	2,292	0.013	0.530	0.004	0.845	-0.018	0.560
Child outcomes							
Denver: Personal-social	2,292	-0.011	0.835	-0.003	0.282	0.001	0.757
Denver: Fine motor	2,292	-0.035	0.621	-0.002	0.520	0.002	0.560
Denver: Language	2,290	-0.035	0.544	-0.005	0.027	0.002	0.456
Denver: Gross motor	2,289	0.022	0.787	-0.002	0.314	-0.001	0.823
Good health	2,292	0.008	0.658	-0.004	0.857	-0.011	0.715

Each row corresponds to a separate regression of a dummy indicating whether the child was surveyed at follow-up or not (=1) on assignment to treatment or control group, the child or family characteristic, and the interaction between these two variables. Standard errors are clustered at the MA level. ‡A reduced 12-item version of the Knowledge of Infant Development Inventory (MacPhee, 1981) was administered. This measure was only collected at BL.

Table B4. Evaluating determinants of effective participation in the PCCSF using randomized assignment and sociodemographic indicators

	>0 visits		>20 visits		Number of visits	
Assignment to treatment and control group	0.108***	0.104***	0.034***	0.033***	1.743***	1.688***
	(0.014)	(0.013)	(0.007)	(0.007)	(0.290)	(0.279)
Age		-0.003***		-0.003***		-0.120***
		(0.001)		(0.001)		(0.033)
Race (white=1)		0.001		0.004		0.167
		(0.011)		(0.006)		(0.274)
Attends a childcare center		-0.038***		-0.007*		-0.433***
		(0.010)		(0.004)		(0.158)
Mother's level of education (completed secondary school or more=1)		0.004		-0.007		-0.143
		(0.012)		(0.006)		(0.281)
Mother works		-0.018		-0.004		-0.328
		(0.011)		(0.007)		(0.309)
Number of symptoms of depression		0.000		0.001		0.028**
		(0.001)		(0.000)		(0.014)
Mother consumed alcohol/drugs during pregnancy		0.005		-0.009		-0.295
		(0.020)		(0.010)		(0.400)
Prenatal checkups (6 or more=1)		0.013		0.009		0.286
		(0.011)		(0.007)		(0.358)
Investments in child (factor)		-0.005		-0.001		-0.039
		(0.005)		(0.003)		(0.100)
Father lives in the home		-0.005		-0.010		-0.431
		(0.011)		(0.008)		(0.331)
Household size		0.001		0.002		0.088
		(0.003)		(0.002)		(0.101)
Per-capita income (>P75=1)		0.001		0.012		0.606*
		(0.013)		(0.008)		(0.339)
Bolsa Familia Beneficiary		0.036***		0.012*		0.500*
		(0.010)		(0.006)		(0.260)
Household is visited by ESF		0.049***		0.017***		0.881***
		(0.011)		(0.006)		(0.245)
Armed violence in neighborhood (last month)		-0.013		-0.010		-0.376
		(0.012)		(0.006)		(0.274)
High perception of health service		0.008		0.004		0.181
		(0.010)		(0.006)		(0.276)
Child development factor		0.001		-0.001		-0.094
		(0.005)		(0.003)		(0.099)
Good health		0.003		0.007		0.416
		(0.010)		(0.006)		(0.255)
Observations	1,961	1,867	1,961	1,867	1,961	1,867
F statistic	58.64	60.33	24.38	24.28	36.08	36.63

Standard errors (clustered) at the MA level in parenthesis. Dependent variable in first row. *** p<0.01, ** p<0.05, * p<0.1

Table B5. Effects of PCCSF on play activities

	ITT	KP F	TOT	N
Play activity factor	-0.026	37.737	-0.016	1,961
	(0.050)		(0.030)	
Reading or looking at pictures (Nb)	-0.046	37.846	-0.028	1,961
	(0.031)		(0.019)	
Telling stories (Nb)	0.031	37.747	0.019	1,961
	(0.030)		(0.018)	
Singing (Nb)	0.013	38.313	0.008	1,961
	(0.030)		(0.018)	
Playing (Nb)	-0.003	37.935	-0.002	1,961
	(0.039)		(0.024)	
Counting or drawing (Nb)	-0.003	37.703	-0.002	1,961
	(0.030)		(0.018)	
Praising them for an specific action (Nb)	-0.016	37.895	-0.010	1,961
	(0.037)		(0.022)	

Play activity factor is computed constructed using CFA of indicator variables of whether any adult did each of the following activities with the child in the last three days: reading or looking at pictures, telling stories, singing, playing, counting or drawing and praising them for a specific action. Nb refers to the number of adults that did that activity with the child in the last three days. Standard errors (clustered) at the MA level in parenthesis. The column ITT displays the results of estimating equation (1) using OLS. The "TOT" column displays the results of estimating a version equivalent to equation (1) using instrumental variables, instrumenting number of total visits with the indicator of random assignment to the treatment and control group. Both models control for: the gender and age of the child, armed violence in neighborhood, whether the household received ESF visits, and the outcome measured at BL (or a variable from the same domain if this was not measured at BL). KP F: Kleibergen-Paap F statistic for weak instruments. *** p<0.01, ** p<0.05, * p<0.1

Table B6. Lee Bounds of intention-to-treat effect on outcomes for which we found a statistically significant effect

	Home environment: Response and discipline	
	Lower	Upper
Coef.	0.094**	0.142***
SE	0.045	0.048
Selected observations	1,960	1,960
%-cut	0.607	0.607
Obs.	2,292	2,292

Lee Bounds were constructed following Lee (2009) procedure. Standard errors are clustered at the MA level. *p<0.10, **p<0.05, ***p<0.01.

Table B7. Effects of PCCSF on Denver subscales

	ITT	TOT		N
		KP F	Effect	
	(1)	(2)	(3)	(4)
Personal-social	-0.061	36.515	-0.037	1,944
	(0.053)		(0.032)	
Fine motor	0.015	36.470	0.009	1,944
	(0.051)		(0.031)	
Language	-0.053	36.416	-0.032	1,942
	(0.051)		(0.032)	
Gross motor	0.043	36.544	0.026	1,941
	(0.051)		(0.031)	

Standard errors (clustered) at the MA level in parenthesis. The column ITT displays the results of estimating equation (1) using OLS. The "TOT" column displays the results of estimating a version equivalent to equation (1) using instrumental variables, instrumenting number of total visits with the indicator of random assignment to the treatment and control group. Both models control for: the gender and age of the child, armed violence in neighborhood, whether the household received ESF visits, and the outcome measured at BL (or a variable from the same domain if this was not measured at BL). KP F: Kleibergen-Paap F statistic for weak instruments. *** p<0.01, ** p<0.05, * p<0.1

Table B8. Effects of PCCSF on the outcomes included in the Home Environment: Responsivity and discipline factor

	ITT	TOT		N
		KP F	Effect	
	(1)	(2)	(3)	(4)
HOME Responsivity factor	-0.030	38.182	-0.018	1,960
	(0.061)		(0.037)	
HOME Acceptance factor	0.083**	37.697	0.051*	1,960
	(0.042)		(0.027)	
HOME Modeling factor	0.141***	37.960	0.086**	1,960
	(0.053)		(0.035)	
Indicator use of negative discipline	0.012	37.981	0.008	1,961
	(0.018)		(0.011)	

Standard errors (clustered) at the MA level in parenthesis. The column ITT displays the results of estimating equation (1) using OLS. The "TOT" column displays the results of estimating a version equivalent to equation (1) using instrumental variables, instrumenting number of total visits with the indicator of random assignment to the treatment and control group. Both models control for: the gender and age of the child, armed violence in neighborhood, whether the household received ESF visits, and the outcome measured at BL (or a variable from the same domain if this was not measured at BL). KP F: Kleibergen-Paap F statistic for weak instruments. We also compute Romano-Wolf (RW) p-values using 1000 repetitions. RW p-value for "Home Acceptance factor" by OLS is 0.140 and by IV is 0.137, and for "Home Modeling factor" by OLS is 0.032 and by IV is 0.043. *** p<0.01, ** p<0.05, * p<0.1

Table B9. Effect of treatment assignment on household likelihood of receiving ESF visits

	Received ESF visits 2016 or after	
Treatment assignment	0.030	0.010
	(0.032)	(0.029)
N	1,961	1,961
Controls	No	Yes

Standard errors (clustered) at the MA level in parenthesis. ESF visits are self-reported. Controls included are the child's gender and age, armed violence in the neighborhood, and whether the household received ESF visits all measured at BL. *** p<0.01, ** p<0.05, * p<0.1

C. Detailed literature review

Luo et al. (2019) evaluate a home visiting program in China for children 6 to 18 months old that aims to improve both child development, by teaching caregivers how to give children psychosocial stimulation, and child health, through structured conversations designed by medical experts and pediatricians. Homes were visited twice a month by health workers who received prior training on the program (five days of classroom instruction and two days of practice in the field). They were also given detailed notes for interactive play-based activities, a kit of early childhood development tools, and a manual with structured conversations on infant and child health topics. For supervision, a strict system of monitoring visits via mobile application was developed, and national-level officials would randomly supervise each worker's visits from time to time. Homes received a total of 24 visits. The authors find immediate positive effects on children's cognitive development, parenting practices, and health and nutrition variables.

Youzafzi et al. (2014) and Gowani et al. (2014) evaluate an enhanced nutritional intervention, a psychosocial stimulation intervention, and a combination of both implemented through 30-minute home visits and 1.5-hour group sessions in rural Pakistan. The interventions are intended for children 2–5 months up to 24 months old and take place monthly. The enhanced nutrition intervention consists of education on nutritional topics and the delivery of a micro-nutrient supplement to 6- to 24-month-old children. The psychosocial stimulation consists of promoting caregiver sensitivity and responsivity through age-appropriate play activities based on the *Care for Child Development* package by UNICEF and the WHO, although it adds many other elements. The visits are carried out by health workers (Lady Health Workers), who are generally women aged 18 to 45 with at least eight years of education and broad acceptance in the community who received two days of training for the nutritional intervention and three days of training for the stimulation intervention, with short on-the-job refresher trainings every month. Youzafzi et al. (2014) finds that the psychosocial intervention had positive effects on child development (cognitive, language, motor, and socio-emotional) 12 months after visiting commenced. The effects persisted in all areas except the socio-emotional domain 24 months after visiting commenced (Youzafzi et al., 2014; Gowani et al., 2014).[24]

Walker et al. (2015) evaluate two interventions: a home visiting program and a center-based intervention, both implemented by the health system, in Kingston, Jamaica. The home visiting intervention is based on a curriculum previously used in Jamaica (Grantham-McGregor, Powell, Walker and Himes, 1991) and adapted to be implemented at scale. CHW guide play sessions between mothers and children to encourage mother-child interactions and show mothers ways to promote their children's development. Children are visited twice a month between the ages of 6 and 18 months. CHW received a three-day training that included reviewing content and methods, practice in small groups, and role-playing. They were also given a handbook with detailed instructions on how to conduct visits and activities. The research team was in charge of supervising the quality of the visits on a monthly basis. The researchers find positive short-term effects on

[24] Youzafzai et al. (2014) also report positive effects of the nutritional intervention on children's cognitive development, but they do not find these interventions (nutritional and psychosocial) to have an additive effect.

children's cognitive development, but not on intermediate variables (parental knowledge or environment quality).

Sylvia et al. (2020) evaluate a home visiting program to improve parenting practices in early childhood in rural areas of one of China's poorest provinces. Households with children18 months old at the beginning of the intervention were visited weekly by cadres from the Family Planning Commission using a curriculum developed by the researchers in tandem with local experts that was loosely related to the Jamaican Reach Up . Cadres were people in charge of enforcing the one child policy. They were re-trained on how to explain the activities to parents and how to motivate them to do them. Six months after the start of the intervention, Sylvia et al. (2020) find positive effects on child development, mediated by higher parental investments in children and improved parenting practices.

Lastly, Blimpo et al. (2016) evaluate a program with two interventions: a basic and intensive version. The basic intervention is a health and nutrition program implemented since 1995 by Gambia's National Nutrition Agency (NaNa) called *Baby Friendly Community Initiative* (BFCI). The program offers education to parents of children 0 to 3 years old and other caregivers through home visiting and group sessions held by previously trained community members. The intensive version (which they call BFCI +) adds a cognitive stimulation component to the services of the basic intervention. The authors find moderate impacts on parental investments and child language development for the intensive version, especially in the region with better socioeconomic conditions, and no effect for the basic intervention.

D. Description of questionnaires and outcome variables

Questionnaires/instruments used for follow-up

Environment quality. We administered 38 items from the Early Childhood HOME Inventory (EC HOME), from the instrument's eight subscales: responsivity, acceptance, learning materials, language stimulation, physical environment, academic stimulation, modeling, and variety (Caldwell and Bradley, 2003). This list of items had been previously administered in an evaluation of a home visiting program in Nicaragua (Lopez Boo, Leer and Kamei, 2020). The HOME Inventory measures home environment quality and focuses on process variables, although it also evaluates certain structural variables. In particular, the responsivity subscale measures a caregiver's responsiveness (emotional and verbal) to a child's needs and whether the relationship is warm. The acceptance subscale evaluates parents' ability to accept and handle children's negative behaviors instead of responding with severe punishments. Learning materials measures the availability of toys, books, and games that promote learning and that are in good condition and accessible to the child. Language stimulation evaluates parents' ability to encourage children's language development through conversation, modeling, and direct teaching. Physical environment evaluates whether the environment is safe for children and whether it is spacious and pleasant. Academic stimulation evaluates parents' direct engagement with children's learning and the acquisition of skills and knowledge considered important during the early years, such as, for example, learning the colors and numbers. Modeling measures parental ability to model desirable and acceptable behaviors, thus transmitting behavioral expectations to children. Finally, the variety

subscale describes a family lifestyle that fosters variety and enriching experiences for the child. The evaluator scores each item by observing interactions (not based on caregiver reports), assigning it a 0 or 1 depending on whether the evaluator observed the behavior during the visit (a few items can be scored through a short interview).

Child development. We administered the Portuguese version of the Denver II Development Screening Test (Frankenburg et al., 1990). The Denver assesses different areas of development (personal-social, fine motor, gross motor, and language) of children from birth until age 6. It consists of 125 items that are scored independently (each item has four possible answers: pass, fail, no answer, no opportunity), and they increase in difficulty as the child grows older. Most of the items are scored through observation, and a few can be evaluated based on what parents report. The final number of items administered depends on the child's age and ability level. Lopez Boo, Cubides Mateus, and Llonch Sabatés (2020) review the psychometric properties of the Portuguese version of this instrument and find the test to be culturally relevant and robust in psychometric terms.

To validate the Denver, we also applied the Bayley-III language scale (referred to as Bayley throughout this study) (Bayley, 2006). The Bayley is one of the most commonly used instruments to measure child development in children under age 42 months (Bayley, 2006).[25] When collecting follow-up measurements, a percentage of the children had outgrown the test's age range. For children whose age allowed us to administer the Bayley, we checked the correlation between this score and the Denver language score, and there is a high correlation (0.83 for receptive language, and 0.77 for expressive language). For these two reasons, we decided not to use the Bayley in the analyses.

Household questionnaire. This instrument was used to collect basic information on the members of the household, including age, level of education, income, and employment; characteristics of the home including number of bedrooms and access to durable goods; and perception of violence in the neighborhood.

Child questionnaire. This instrument gathers information about the child's health status and childcare attendance; the availability of learning materials in the home; simple activities—like reading, singing, and counting numbers—that adults do with the child; and details about the mother, father, and primary caregiver.

The socio-demographic and child-related characteristics were gathered by professional interviewers. The other instruments were administered by observers: students, or professionals in the fields of psychology, education, social work, or similar areas. To collect the information, interviewers and observers were paired up in two-person teams. They arrived together and both remained at the home for the entire interview.

[25] It consists of five scales: motor, cognitive, language, socio-emotional, and adaptive behaviors. In this evaluation, we only administered the language scale, which evaluates both receptive and expressive language. The first subscale consists of 49 items, and the second 48 items, which increase in difficulty with the child's age. Each item can be scored 1 or 0 based on whether the child was able to do the activity. The final number of items administered depends on the child's age and ability level. Unlike the Denver Test, the Bayley is a diagnostic rather than screening instrument, so it can give more information.

Before beginning fieldwork, the interviewers and observers were trained. Interviewers' training lasted 6 days: 4 focusing on theory and 2 on practice. Training for observers lasted 9 days, with 7 focusing on theory and 2 focusing on practice. The survey firm was in charge of providing training on the questionnaires and administering the HOME. Training on the Denver was provided by a doctor specializing in maternal and newborn health and research on health issues in high-risk infants and child development. Additionally, a three-day pilot program was carried out in a Fortaleza neighborhood with socioeconomic conditions similar to those of Region V.

Outcome variables

Based on the theory of change of the Program, we defined the set of variables for which we would estimate the effect of the PCCSF:

- Factor: Cognitive stimulation. Based on the score on the learning materials, language stimulation, learning stimulation, and variety subscales of the HOME, we constructed a factor and subtracted its mean (of the control group) and divided by its standard deviation (of the control group).
- Factor: Responsivity and behavioral strategies. Using the score from the responsivity, acceptance, and modeling subscales of the HOME Inventory and the indicator for the use of negative discipline, we constructed a factor and subtracted its mean (of the control group) and divided by its standard deviation (of the control group).[26] The use of negative discipline indicator was constructed based on a set of disciplinary strategies, and primary caregivers were asked whether they used these strategies when the child behaved poorly. More specifically, they were asked whether they used: 1) reprimands, scolding, or yelling; 2) slapping or smacking; 3) hitting the child with their hands or fist; 4) hitting them with a band, belt, rod, or other object; 5) shutting them in a room or isolating them; 6) leaving the child outside the home; 7) shaking the child. If the primary caregiver responded yes to any of the seven strategies listed above, the household was considered to use negative discipline. Otherwise, it was assumed that negative discipline is not used.
- Factor: Investments in child. We constructed a factor using the variables that evaluate parental investment in children. More specifically, we included:
    - Hours the primary caregiver spent caring for the child on a typical weekday. The primary caregiver was asked how much time on average they spent caring for the child on each weekday, thinking back on the last 30 days. The average number of hours reported for Monday through Friday was calculated, and min-max normalization was applied to this value.
    - Hours the primary caregiver spent caring for the child on a typical weekend day. The primary caregiver was asked how much time on average they spent caring for the child on each weekend day, thinking back on the last 30 days. The average number of hours reported for Saturday and Sunday was calculated, and min-max normalization was applied to this value.
    - Hours the child spent watching television on a typical weekday. The primary caregiver was asked how many hours on average the child spends watching

[26] To construct the factor, the negative discipline indicator is first reversed.

television, whether in the home or elsewhere, on a typical weekday. This value was reversed and min-max normalization was applied to it.
    - Hours the child spent watching television on a typical weekend day. The primary caregiver was asked how many hours on average the child spends watching television, whether in the home or elsewhere, on a typical weekend day. This value was reversed and min-max normalization was applied to it.
    - Indicators for activities with the child. The primary caregiver was asked whether any adult had done any of the following activities with the child in the last three days: 1) reading or looking at pictures in books, magazines, or newspapers; 2) telling stories; 3) singing; 4) playing; 5) counting or drawing; and 6) praising them for a specific action.

  The factor is subtracted by its mean (of the control group) and divided by its standard deviation (of the control group).
- Child development factor. We performed an exploratory factor analysis using the score from the personal-social, fine motor, and language subscales of the Denver. We use the continuous scoring in each subscale internally normalized for age—using nonparametric regressions—with a mean zero and standard deviation one in the full sample. The analysis yields a single factor with an eigenvalue greater than 1 and includes the three subscales (all with loadings greater than 0.4). The factor is then subtracted by its mean (of the control group) and divided by its standard deviation (of the control group).
- Child's good health. The primary caregiver was asked about the child's health status. More specifically, they were asked: 1) whether the child had any health issue, ailment, or accident in the past 30 days, 2) whether the child had diarrhea in the past 30 days, and 3) whether the child had any illness accompanied by a high fever in the past 30 days. The child is considered to be in good health if the caregiver answered "no" to these three questions.

Table D1. Bivariate correlations between intermediate and final variables

	Child development factor	Good health
Home environment: Cognitive stimulation	0.222***	0.021
Home environment: Response and discipline	0.038*	0.044*
Investments in child (factor)	0.112***	-0.062***

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

Table D2. Socioeconomic gradients of variables measured at follow-up

	N	Home environment		Investments in child (factor)	Child development factor	Good health
		Cognitive stimulation	Response and discipline			
By household income level						
First quartile (Q1)	491	-0.172	0.125	-0.097	-0.110	0.385
		(1.129)	(0.904)	(1.043)	(1.128)	(0.487)
Second quartile	493	-0.062	0.098	-0.109	0.022	0.418
		(0.946)	(0.885)	(0.997)	(0.799)	(0.494)
Third quartile	487	0.004	0.046	-0.028	-0.028	0.421
		(0.944)	(0.979)	(1.020)	(0.969)	(0.494)
Fourth quartile (Q4)	490	0.236	-0.016	0.191	0.086	0.445
		(0.830)	(0.935)	(0.946)	(0.941)	(0.497)
Q1=Q4 (p-value)		**0.000**	**0.017**	**0.000**	**0.003**	**0.057**
By mother's level of education						
Primary school incomplete or less (M1)	499	-0.394	0.056	-0.253	-0.072	0.379
		(1.220)	(0.913)	(1.056)	(0.851)	(0.486)
Primary school complete or secondary school incomplete	721	0.016	0.089	-0.017	-0.076	0.434
		(0.878)	(0.873)	(1.022)	(1.050)	(0.496)
Secondary school complete or higher (M3)	709	0.270	0.033	0.166	0.100	0.426
		(0.766)	(0.993)	(0.921)	(0.963)	(0.495)
M1=M3 (p-value)		**0.000**	0.681	**0.000**	**0.001**	0.100

Household income and mother's level of education measured at follow-up. The table shows the mean (and standard deviation in parentheses) of each variable and for each group. P-value of the difference in means test. P-values < 0.10 in bold.

References

Ajzenman, N., Bertoni, E., Elacqua, G., Marotta, L., and Méndez Vargas, C. 2020. "Altruism or Money? Reducing Teacher Sorting using Behavioral Strategies in Perú." IDB Working Paper Series, IDB-wp-01143.

Andrew, A., Attanasio, O., Augsburg, B., Day, M., Grantham-McGregor, S., Meghir, C., Mehrin, F., Pahwa, S., and Rubio-Codina, M. 2020. "Effects of a scalable home-visiting intervention on child development in slums of urban India: evidence from a randomized controlled trial." *Journal of Child Psychology and Psychiatry*, 61(6), 644-652.

Ara G, Khanam M, Papri N, et al. 2018. "Peer counselling improves breastfeeding practices: A cluster randomized controlled trial in urban Bangladesh." *Matern Child Nutr*., 14:e12605.

Araujo, M.C. Dormal M., Rubio M., and Schady N. 2019. Cuna Mas Home Visiting at Scale and Child Development, IDB mimeo.

Attanasio, O.P., Fernández, C., Fitzsimons, E.O., Grantham-McGregor, S.M., Meghir, C., and Rubio-Codina, M. 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." *BMJ (online)*, 349 (September), 1-12.

Baker-Henningham, H., Powell, C., Walker, S., and Grantham-McGregor, S. 2003. "Mothers of undernourished Jamaican children have poorer psychosocial functioning and this is associated with stimulation provided in the home." *European Journal of Clinical Nutrition*, 57(6), 786.

Baker-Henningham, H., and Lopez Boo, F. 2010. "Early Childhood Stimulation Interventions in Developing Countries: A comprehensive literature review." Washington, D.C.: IDB

Bayley, N. 2006. *Bayley Scales of Infant and Toddler Development: Bayley III*. United States: Harcourt Assessment, Psych. Corporation.

Berlinsky, S. and Schady, N. (eds). 2015. Los primeros años: El bienestar infantil el papel de las políticas públicas. Development in the Americas (DIA). Washington, D.C.: IDB.

Bhalotra, S., Rocha, R., and Soares, R.R. 2016. "Does universalization of health work? Evidence from health systems restructuring and maternal and child health in Brazil." Institute for Social and Economic Research, Working Paper No. 2016-16.

Black, M. M. and Dewey, K. G. 2014. "Promoting equity through integrated early child development and nutrition interventions." *Ann. NY Acad. Sci.* 1308, 1–10

Black, M. M., et al. 2017. "Early childhood development coming of age: science through the life course." *Lancet*, 389(10064), 77-90.

Blimpo, M., Carneiro, P., Jarvis, P. and Lahire, N. 2016. "Improving Parental Investments in Children: Experimental Evidence from the Gambia." IZA Conference, presentation

Brenner, J.L.; Kabakyenga, J.; Kyomuhangi, T.; Wotton, K.A.; Pim, C.; Ntaro, M.; Bagenda, F.N.; Gad, N.R.; Godel, J.; Kayizzi, J.; et al. 2011. "Can volunteer community health workers decrease child morbidity and mortality in southwestern Uganda? An impact evaluation." PloS ONE, 6

Caldwell, B. M. 1967. "Descriptive Evaluations of Child Development and of Developmental Settings." *Pediatrics,* 40(1), 46-54.

Caldwell, B.M. and Bradley, R.H. 2003. Home Observation for Measurement of the Environment: Administration Manual. Tempe, AZ: Family & Human Dynamics Research Institute, Arizona State University.

Chan, M. 2013. "Linking child survival and child development for health, equity, and sustainable development." *Lancet,* 381, 1514-1515.

Cheng, L.Y., Wang, X., and Mo, P.K. 2019. "The effect of home-based intervention with professional support on promoting breastfeeding: a systematic review." *Int J Public Health* 64, 999–1014.

Digirolamo, A.M., Stansbery, P, and Lung'aho, M. 2014. "Advantages and challenges of integration: opportunities for integrating early childhood development and nutrition programming." *Ann N Y Acad Sci*; 1308: 46-53.

Frankenburg, W. K., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelmann, N., and Shapiro, H. 1990. *The Denver II technical manual*. Denver, CO: Denver Developmental Materials.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S.M., and Grantham-McGregor, S. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science*, 344 (6187), 998-1001.

Gowani, S., Yousafzai, A.K., Armstrong, R., and Bhutta, Z.A. 2014. "Cost effectiveness of responsive stimulation and nutrition interventions on early child development outcomes in Pakistan." *Ann. N.Y. Acad. Sci*. 1308 (1), 149–161.

Grantham-McGregor, S.M., Powell, C.A., Walker, S.P., and Himes, J.H. 1991. "Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study." *The Lancet*, 338(8758), 1-5.

Grantham-McGregor, S., Powell, C., Walker, S., Chang, S., and Fletcher, P. 1994. "The long-term follow-up of severely malnourished children who participated in an intervention program." *Child development*, 65(2), 428-439.

Hamadani, J.D., Huda, S.N., Khatun, F., and Grantham-McGregor, S.M. 2006. "Psychosocial Stimulation Improves the Development of Undernourished Children in Rural Bangladesh." *The Journal of Nutrition*, 136(10), 2645-2652.

Hamadani, J.D., Mehrin, S.F., Tofail, F., Hasan, M.I., Huda, S.N., Baker-Henningham, H., Grantham-McGregor, S.M. 2019. "Integrating an early childhood development programme into Bangladeshi primary health-care services: an open-label, cluster randomized controlled trial." *Lancet Glob. Health*, 7(3), e 366–e375.

Heckman, J.J. 2008. "School, Skills, and Synapses." NBER Working Paper No. 14064.

Hone, T., Rasella, D., Barreto, M.L., Majeed, A., and Millett, C. 2017. "Association between expansion of primary healthcare and racial inequalities in mortality amenable to primary care in Brazil: A national longitudinal analysis." *PLOS Medicine*, 14(5): e1002306.

IBGE. (n.d.). População. Retrieved May 19, 2020 from https://cidades.ibge.gov.br/brasil/ce/fortaleza/panorama

Inter-American Development Bank. 2019. Estudio de grupos focales sobre la implementación del Programa Cresça com Seu Filho conducido por Juliana Freitas. Mimeo

Ehrlich, I., and Yong, Y. 2013. "Equilibrium Health Spending and Population Aging in a Model of Endogenous Growth. Will the GDP Share of Health Keep Rising?" *Journal Of Human Capital*, 4(7), 411-447.

Janmohamed, A., Sohani, N., Lassi, Z.S., and Bhutta, Z.A. 2020. "The Effects of Community Home Visit and Peer Group Nutrition Intervention Delivery Platforms on Nutrition Outcomes in Low and Middle-Income Countries: A Systematic Review and Meta-Analysis." *Nutrients*, *12*, 440.

J-PAL. 2020. Encouraging early childhood stimulation from parents and caregivers to improve child development. Retrieved May 12, 2020 from https://www.povertyactionlab.org/policy-insight/encouraging-early-childhood-stimulation-parents-and-caregivers-improve-child

Lee, D. S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects*". Review of Economic Studies*, 76(3), 1071–1102.

Leer, J., and Lopez Boo, F. 2018. "Assessing the quality of home visit parenting programs in Latin America and the Caribbean." *Early Child Development and Care*, 189(13), 2183-2196

List, J., Suskind, D., and Supplee, L.H. (Eds.). 2021. *The Scale-Up Effect in Early Childhood and Public Policy* (1st ed.). No. ISBN 9780367422479. Routledge.

Lopez Boo, F. and Encina, J. 2015. Informe de capacitaciones de Reach Up en Fortaleza, mimeo.

Lopez Boo, F., Leer, J., and Kamei, A. 2020. "Can community monitoring improve the effect of home visits on child development? Experimental evidence from Nicaragua." Mimeo

Lopez Boo, F., Cubides Mateus, M., and Llonch Sabatés, A. 2020. "Initial psychometric properties of the Denver II in a sample from Northeast Brazil." *Infant Behavior and Development*, 58.

Lozoff, B., Smith, J.B., Clark, K.M., Perales, C.G., Rivera, F., and Castillo, M. 2010. "Home intervention improves cognitive and social-emotional scores in iron-deficient anemic infants." *Pediatrics*, 126(4), e884-e894.

Luo R., Emmers, D., Warrinnier B., Rozellec, S. and Sylvia, S. 2019. "Using community health workers to deliver a scalable integrated parenting program in rural China: A cluster-randomized controlled trial." *Social Science & Medicine*, 239: 112545.

MacPhee, D. 1981. Knowledge of Infant Development Inventory. Survey of child care experiences-Manual.

McCormick, M.C, et al. 2006. « Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age for the Infant Health and Development Program." *Pediatrics*, 117(3), 771-780.

Nahar, B., Hamadani, J.D., Ahmed, T., Tofail, F., Rahman, A., Huda, S.N., and Grantham-McGregor, S.M. 2009. "Effects of a psychosocial stimulation on growth and development of

severely malnourished children in a nutrition unit in Bangladesh." *European Journal of Clinical Nutrition*, 63(6), 725-731.

Nair, N., Tripathy, P., Sachdev, H.S., Pradhan, H., Bhattacharyya, S., Gope, R., Gagrai, S., Rath, S., Rath, S., Sinha, R., et al. 2017. "Effect of participatory women's groups and counselling through home visits on children's linear growth in rural eastern India: A cluster-randomised controlled trial." *Lancet Glob. Health*, 5.

Núcleo de Ciencia para la Infancia-NCI. 2019. "Impactos da Estratégia Saúde da Família e Desafios para o Desenvolvimento Infantil." Estudio 5. Retrieved from https://ncpi.org.br/wp-content/uploads/2019/12/AF_NCPI-WP_n5_2019_online_v2.pdf

Olds, D.L., Henderson, C.R., Kitzman, H., Eckenrode, J., Cole, R., and Tatelbaum, R. 1999. "Prenatal and infancy home visitation by nurses: recent findings." *Future of children*, 9(1), 44-65.

Olds, D.L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D.W., …, Holmberg, J. 2004. "Effects of Nurse Home-Visiting on Maternal Life Course and Child Development: Age 6 Follow-Up Results of a Randomized Trial." *Pediatrics*, 114(6), 1550-1559.

Powell, C., and Grantham-McGregor, S. 1989. "Home Visiting of Varying Frequency and Child Development." *Pediatrics*, 84(1), 157-164.

Radloff, L.S. 1977. "The CES-D scale: A self-report depression scale for research in the general population." *Applied psychological measurement*, 1(3), 385-401.

Ribeiro, F.G., Braun, G., Carrero, A., da Silva Teixeira, G., and Petrucci Gigante, D. 2018. "An empirical assessment of the Healthy Early Childhood Program in Rio Grande do Sul State, Brazil." *Cad. Saúde Pública*, 34(4).

Richter, L.M., Daelmans, B. 2017. "Investing in the foundation of sustainable development: pathways to scale up for early childhood development." *Lancet,* 389(10064), 103-118.

Sama-Miller, E., Akers, L., Mraz-Esposito, A., Zukiewicz, M., Avellar, S., Paulsell, D., and Grosso, P.D. 2017. Home visiting evidence of effectiveness review: Executive summary (Tech. Rep.). *Mathematica Policy Research*

Shah, R., Kennedy, S., Clark, M.D., et al. 2016. "Primary Care-Based Interventions to Promote Positive Parenting Behaviors: A Meta-analysis." *Pediatrics*, 137(5).

Shakya, P., Kunieda, M. K., Koyama, M., Rai, S. S., Miyaguchi, M., Dhakal, S., … Jimba, M. 2017. "Effectiveness of community-based peer support for mothers to improve their breastfeeding practices: A systematic review and meta-analysis." *PLoS One*, 12(5), e0177434.

Shi, L., Zhang, J., Wang, Y., Caulfield, L.E., Guyer, B. 2010. "Effectiveness of an educational intervention on complementary feeding practices and growth in rural China: A cluster-randomised controlled trial." *Public Health Nutr*., 13, 556–565.

Supplee, L.H., and Duggan, A. 2019. "Innovative research methods to advance precision in home visiting for more efficient and effective programs." *Child Development Perspective*, 13(3), 173-179.

Sylvia, S., Warrinnier, N., Luo, R., Yue, A., Attanasio, O., Medina, A., and Rozelle, S. 2020. "From Quantity to Quality: Delivering a Home-based Parenting Intervention through China's Family Planning Cadres." *The Economic Journal*, ueaa114, https://doi.org/10.1093/ej/ueaa114

Verch, K. 2017. Primeira Infância Melhor: Transforming the Attention towards the First Years of Life in Latin America: Challenges and Achievements of a Public Policy in Southern Brazil. Monograph, IDB.

Walker, S. P., Baker-Henningham, H., Chang, S.M., Powell, C.A., Lopez Boo, F., and Grantham-McGregor, S. 2017. "Implementation of parenting interventions through health services in Jamaica." *Vulnerable Children and Youth Studies*, 13(2), 127-141.

Walker, S.P., Chang, S.M., Vera-Hernandez, M., and Grantham-McGregor, S. 2011. "Early childhood stimulation benefits adult competence and reduces violent behavior." *Pediatrics*, 127(5), 849-857.

Walker, S.P., Powell, C., Chang, S.M., Baker-Henningham, H., Grantham-McGregor, S., Vera-Hernandez, M., and Lopez Boo, F. 2015. "Delivering Parenting Interventions through Health Services in the Caribbean." IDB Working Paper Series, IDB-WP-642.

Worku, B. N., Abessa, T. G., Wondafrash, M., Lemmens, J., Valy, J., Bruckers, L., Kolsteren, P., and Granitzer, M. 2018. "Effects of home-based play-assisted stimulation on development performances of children living in extreme poverty: a randomized single-blind controlled trial." *BMC pediatrics*, 18(29).

Yousafzai, A.K., Rasheed, M.A., Rizvi, A., Armstrong, R., Bhutta, Z.A., 2014. "Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial." *The Lancet*, 384 (9950), 1282–1293.

Figure 1. Randomization process

[Figure1]

Figure 2. Number of visits a CHW conducted in a month

[Figure2]

Table 1. Number of complete PCCSF visits the child received

Number of visits	%
0	88.99
1-5	3.29
6-10	2.67
11-15	1.13
16-20	0.51
21-25	0.93
26-30	1.03
>30	1.44

The percentage is calculated based on the number of children in treatment MA.

Table 2. Baseline variable comparisons

Socio-demographic indicators in BL	N	Control (SD)	Treatment (SD)	Mean difference (SE)
Characteristics of child:				
Age (in months)	1,961	18.910	18.745	-0,165
		(5.930)	(5.749)	(0.283)
Gender (male=1)	1,961	0.524	0.505	-0.019
		(0.500)	(0.500)	(0.023)
Race (white=1)	1,961	0.299	0.281	-0.018
		(0.458)	(0.450)	(0.021)
Attends a childcare center	1,961	0.098	0.091	-0.008
		(0.298)	(0.287)	(0.014)
Characteristics of mother				
Years of education	1,603	9.342	9.597	0.255
		(2.880)	(2.805)	(0.157)
Works	1,870	0.371	0.388	0.017
		(0.483)	(0.488)	(0.025)
Number of symptoms of depression†	1,870	8.796	8.536	-0.260
		(7.665)	(7.586)	(0.396)
Environment quality and investment in child				
HOME: Responsivity	1,961	1.331	1.399	0.069
		(1.406)	(1.487)	(0.085)
HOME: Acceptance	1,961	4.582	4.539	-0.043
		(0.879)	(0.956)	(0.045)
Investments in child (factor)	1,961	-0.019	-0.023	-0.004
		(0.977)	(0.947)	(0.049)
Children's books in the home	1,938	1.283	1.224	-0.059
		(3.504)	(3.025)	(0.148)
Characteristics of household				
Father lives in the home	1,961	0.689	0.671	-0.018
		(0.463)	(0.470)	(0.022)
Per-capita income (R$)	1,915	259.088	255.422	-3.665
		(218.997)	(192.731)	(10.432)
Bolsa Familia Beneficiary	1,961	0.788	0.790	0.002
		(0.409)	(0.407)	(0.021)
Armed violence in neighborhood (in last month)	1,961	0.507	0.448	-0.059*
		(0.500)	(0.497)	(0.030)
Household is visited by ESF	1,961	0.480	0.546	0.066*
		(0.500)	(0.498)	(0.034)
Child Outcomes				
Child development factor	1,959	0.017	0.034	0.018
		(0.961)	(0.982)	(0.046)
Denver: Personal-social	1,961	14.712	14.669	-0.043
		(4.077)	(3.984)	(0.185)
Denver: Fine motor	1,961	16.892	16.870	-0.021
		(3.260)	(3.333)	(0.158)
Denver: Language	1,959	17.901	17.929	0.028
		(4.696)	(4.951)	(0.227)
Good health	1,961	0.367	0.363	-0.004
		(0.482)	(0.481)	(0.023)

OLS estimation with clustered standard errors by MA. †Measured using the CES-D self report scale, 12 items (Radloff, 1977). If the person's score was ≥ 10, he or she is considered to have significant symptoms of depression. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Effects of PCCSF on final and intermediate outcomes

	ITT	TOT		N
		KP F	Effect	
	(1)	(2)	(3)	(4)
		Final outcomes		
Child development factor	-0.032	36.435	-0.019	1,942
	(0.046)		(0.028)	
Good health	-0.014	37.867	-0.009	1,961
	(0.023)		(0.014)	
		Intermediate outcomes		
Home environment: Cognitive stimulation	0.004	37.960	0.003	1,960
	(0.049)		(0.030)	
Home environment: Responsivity and discipline	0.120**	37.960	0.073**	1,960
	(0.048)		(0.031)	
Investments in child (factor)	-0.027	37.729	-0.017	1,961
	(0.052)		(0.032)	

Standard errors (clustered) at the MA level in parenthesis. The column ITT displays the results of estimating equation (1) using OLS. The "TOT" column displays the results of estimating a version equivalent to equation (1) using instrumental variables, instrumenting number of total visits with the indicator of random assignment to the treatment and control group. Both models control for: the gender and age of the child, armed violence in neighborhood, whether the household received ESF visits, and the outcome measured at BL (or a variable from the same domain if this was not measured at BL). KP F: Kleibergen-Paap F statistic for weak instruments. The variables were grouped into blocks by type of outcomes to calculate Romano-Wolf (RW) p-values which account for multiple hypothesis testing. More specially, the variables of the first panel are one block, and those of the second are another. The RW p-value of the effect on the variable "Home environment: Response and discipline" by OLS is 0.036 and by IV is 0.036. 1000 repetitions. *** p<0.01, ** p<0.05, * p<0.1

As a robustness exercise, we estimated equation (1) using OLS (i) without including controls, (ii) including the controls listed minus the baseline measure for child development, and (iii) including all controls plus fixed effects of UAPS. The results were robust regardless of the specification. We also estimated the effects of receiving at least one visit and receiving more than 20 visits. The magnitude of the coefficients changes relative to the results in column (3), but the significance of the effects remains the same.