Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers | Cowles Foundation

2-1-2015

Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia

Orazio P. Attanasio

Sarah Cattan

Emla Fitzsimons

Costas Meghir

Marta Rubio-Codina

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

 Part of the Economics Commons

Recommended Citation

Attanasio, Orazio P.; Cattan, Sarah; Fitzsimons, Emla; Meghir, Costas; and Rubio-Codina, Marta, "Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia" (2015). *Cowles Foundation Discussion Papers*. 2415.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/2415

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

ESTIMATING THE PRODUCTION FUNCTION FOR HUMAN CAPITAL: RESULTS FROM A RANDOMIZED CONTROLLED TRIAL IN COLOMBIA

By

Orazio Attanasio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina

February 2015

Revised July 2018

COWLES FOUNDATION DISCUSSION PAPER NO. 1987R2



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia

Orazio Attanasio, Sarah Cattan, Emla Fitzsimons,
Costas Meghir, and Marta Rubio-Codina*

July 4, 2018

Abstract

We examine the channels through which a randomized early childhood intervention in Colombia led to significant gains in cognitive and socio-emotional skills among a sample of disadvantaged children aged 12 to 24 months at baseline. We estimate the determinants of parents' material and time investments in these children and evaluate the impact of the treatment on such investments. We then estimate the production functions for cognitive and socio-emotional skills. The effects of the program can be explained by increases in parental investments, emphasizing the importance of parenting interventions at an early age.

*Attanasio: University College London and Institute for Fiscal Studies (o.attanasio@ucl.ac.uk). Cattan: Institute for Fiscal Studies (sarah_c@ifs.org.uk). Fitzsimons: UCL Institute of Education and Institute for Fiscal Studies (e.fitzsimons@ucl.ac.uk). Meghir: Yale University, NBER, IZA and Institute for Fiscal Studies (c.meghir@yale.edu). Rubio-Codina: Inter-American Development Bank and Institute for Fiscal Studies (martarubio@iadb.org).

Acknowledgements: We are grateful to three anonymous referees and the editor for detailed comments. We also thank participants at the NBER Summer Institute, Barcelona GSE Summer Forum and Montreal CIREQ Applied Microeconomics on Fertility and Child Development and seminars at Stanford University, University of Chicago, Oxford University, Cornell University, Bristol University and the Institute for Fiscal Studies for their comments. We are grateful to the Economic and Social Research Council (Grant ES/G015953/1), the Inter-American Development Bank, the International Growth Centre, and the World Bank for funding the intervention and data collection. Some of this research was financed by the European Research Council's Advanced Grant 249612 and by the Grand Challenges Canada Prime Award 0072-03 (sub-award reference number 560450). Sarah Cattan gratefully acknowledges financial assistance from the British Academy Postdoctoral Fellowship pf140104, as well as from the European Research Council's Grant Agreement No. 240910. We thank the NIH for funding under grant R01 HD072120 as well as the Cowles foundation and the ISPS at Yale for financial assistance. All errors are the responsibility of the authors.

1 Introduction

The first five years of life lay the basis for lifelong outcomes (Almond and Currie, 2011). Due to rapid brain development and its malleability during the early years (Knudsen, 2004; Knudsen et al., 2006), investments during this period play a crucial role in the process of human capital accumulation. At this time however, many children are exposed to risk factors such as poverty, malnutrition and non-stimulating home environments preventing them from reaching their full potential, particularly in developing countries (Black et al., 2016; Grantham-McGregor et al., 2007; Lu et al., 2016). Thus children from poor backgrounds accumulate developmental deficits from a very early age (Lancet, 2016; Rubio-Codina et al., 2015). These factors are likely to play an important role in the intergenerational transmission of poverty.

There is increasing evidence that early childhood interventions can help overcome these detrimental factors and have positive effects on children's development in both the short- and long-term. Examples include the Jamaica study (Grantham-McGregor et al. (1991), Walker et al. (2011) and Gertler et al. (2014)), the Perry Preschool program (Heckman et al., 2010) and the Abecedarian experiment (Campbell and Ramey (1994), Campbell et al. (2014)). In Attanasio et al. (2014), we present the impacts of an 18-month long early childhood intervention in Colombia targeted at disadvantaged children aged 12-24 months old at baseline and evaluated by a randomized controlled trial. The intervention was based on the Jamaican model in that it offered psychosocial stimulation via weekly home visits and micronutrient supplementation. However, unlike the Jamaican program, it was designed to be scalable by training local women involved in the implementation of a large welfare program to administer the weekly home visits.

The randomized controlled trial (RCT) used to evaluate each arm of the intervention showed that stimulation led to highly significant improvements in cognition and

language development measured immediately following the end of the intervention.[1] Micronutrient supplementation did not affect any outcome observed in the data.

Building on these results, the main aim of this paper is to understand *how* the stimulation component of the intervention led to improvements in child development. For example, it could have led parents to make greater material and time investments in their children. But it could also have changed the production function for child skills, through the direct effect of the home visits as a new input or by changing the effectiveness of parental inputs. In what follows, we build a model of parental investments and child skill formation to tease out the relative importance of these different mechanisms, a crucial step to better focus and increase the sustainability of interventions in the future.

We start by estimating the determinants of parental investments and assessing how the intervention changed parental choices. Indeed the way parents respond to such programs, which can be seen as a type of in-kind transfer, is an open question: the intervention could lead parents to reinforce their engagement with the child or instead crowd-out their investments. Gelber and Isen (2010), for example, provide evidence that the US early childhood program Head Start led to an increase in parental involvement, thus crowding-in household resources. In our treatment of the question here, we exploit the experimental variation induced by the RCT and distinguish between material investments (e.g. books and toys around the house) and time investments (e.g. time spent by an adult in the household on education activities with the child).[2]

We then estimate production functions for child cognitive and socio-emotional skills. The main inputs we specify are baseline child skills, maternal skills, and material and quality time investments, which we treat as endogenous. Within this framework, we

[1] Cognition improved by 26% of a Standard Deviation (SD) (p-value 0.002) and receptive language by 22% of a SD (p-value 0.032). These reported p-values are adjusted for testing 12 hypotheses.

[2] See DelBoca et al. (2014) for a structural model of household choices and child development based on the PSID Child Development Supplement data and also including time and resource investments.

quantify by how much changes in parental investments contributed to improving child outcomes in the treatment group. We also test whether the intervention changed the parameters of the production function, which, as discussed above, could reflect the direct effect of the stimulation provided by the home visitors or a change in the productivity of inputs.

The two waves of data we use were collected just before and just after the intervention and contain rich measures of child development, maternal skills and parental investments. Importantly, we collect information on materials and activities that have an educational aspect, thus enabling a clear interpretation of parental behavior as investments in their children. To our knowledge, our sample is one of the largest ever collected with this type of data in the literature evaluating stimulation programs. Even with such rich data however, estimating the parameters governing the skill formation process remains challenging for two reasons. First, inputs and outputs are likely to be measured with error. Second, inputs, especially investments, can be endogenous, if parental decisions respond to shocks or inputs that are unobserved to the econometrician. To deal with the measurement error issue, we use dynamic latent factor models as in Cunha and Heckman (2008) and Cunha, Heckman, and Schennach (2010). To deal with the possible econometric endogeneity of investments, we use an instrumental variable strategy.

The estimates of the investment functions reveal important information about some of the drivers of developmental inequality: children with better initial cognitive skills receive more investments and, crucially, mothers with higher skill levels invest more in their children *given* the child's skills. In line with the existing literature, we find that a child's current stock of skills fosters the development of future skills, although we do find mean reversion.[3] Second, parental investments and in particular our measure

[3]These features of the technology of skill formation are often referred to as *self-productivity* and *cross-productivity* (Cunha et al., 2006).

of material investments are an important determinant of future cognitive and socio-emotional skills. This becomes even more evident once we control for the endogeneity of such investments in line with results from Cunha, Heckman, and Schennach (2010) and Attanasio et al. (2017) in very different contexts.[4]

With respect to the mechanisms underlying the impacts, we find that the intervention significantly increased parental investments among treated families compared to non-treated ones. This increase is the dominant reason underlying the observed impacts. The direct effect of the intervention, instead, is both small and insignificant. These two findings mean that the gains in cognitive and socio-emotional skills among children who received the intervention are mainly explained by changes in parental investments and imply that having the home visitor merely interact with the child for an hour a week, without trying to strengthen parenting practices, would have been unlikely to benefit children. This, together with our mean reversion result, emphasizes the key importance of improving parenting practices for the success and the longer term sustainability of early childhood interventions.

Along with Heckman, Pinto, and Savelyev (2013) and a few other papers (Attanasio, Meghir, and Santiago, 2012; Duflo, Hanna, and Ryan, 2012; Todd and Wolpin, 2006), our paper illustrates how data from randomized trials can be profitably combined with behavioral models to go beyond the estimation of experimentally induced treatment effects and interpret the mechanisms underlying them. While there is a large literature evaluating the impact of early childhood interventions on child development,

[4]The former use the children of the National Longitudinal Survey of Youth 1979, a longitudinal panel following the children of a representative sample of women born between 1956 and 1964 in the US. The latter use the Young Lives Survey for India, a longitudinal survey following the lives of children in two age-groups: a Younger Cohort of 2,000 children who were aged between 6 and 18 months when Round 1 of the survey was carried out in 2002, and an Older Cohort of 1,000 children then aged between 7.5 and 8.5 years. The survey was carried out again in late 2006 and in 2009 (when the younger children were about 8, the same age as the Older Cohort when the research started in 2002). See also Helmers and Patnam (2011) for the estimation of a linear production function in India. Finally, and also in line with the existing literature, we find that current skills and parental investments are complementary in the production of future skills, meaning that returns to investments are higher for children with better initial conditions.

our paper innovates by complementing the information obtained from the RCT of a specific intervention with a model of skill formation and parental investment in order to understand the mechanisms behind the observed impacts.

In this sense, our paper shares the motivation of Heckman, Pinto, and Savelyev (2013) who document the channels through which the Perry Pre-School Program produced gains in adult outcomes. But our focus and methodology are different: Heckman, Pinto, and Savelyev (2013) perform a mediation analysis that decomposes linearly the treatment effects on adult outcomes into components attributable to early changes in different personality traits. Instead, we use a model in which parents make investment choices and human capital accumulates according to a production function, so as to interpret and explain the impacts induced by a successful intervention.

The focus of our intervention is also different. Unlike the Jamaican intervention, which targeted malnourished children, and the Perry Preschool Program, which targeted children with specifically low cognition, we target a broader population. Our subjects are drawn from the beneficiaries of the Colombian Conditional Cash Transfer (CCT) program *Familias en Acción*, which covers the poorest 20% of the population.[5] In this sense, our program has the potential to serve as a model for early childhood policy that could be broadly implemented alongside CCT programs or other welfare programs targeting poor families.

The paper proceeds as follows. Section 2 provides some background on the intervention. Section 3 describes the data and the factor model approach we take to extract error-free measures of children's skills, parental skills and investments. Section 4 discusses the short-term impacts of the intervention and some suggestive evidence of its underlying mechanisms. Section 5 presents our theoretical framework and its empirical implementation. Section 6 presents the estimates of the model and discusses their implications for our understanding of the intervention. Section 7 concludes.

[5]See Attanasio et al. (2010) for a description and evaluation of that program.

2 Background on the intervention and its evaluation

The early childhood program analyzed in this paper was targeted at children aged between 12 and 24 months living in families receiving the Colombian CCT program, which targets the poorest 20% of households in the country. The intervention lasted 18 months, starting in early 2010. Appendix A contains a detailed description of the program's design, implementation and delivery. Here we summarize the key aspects.

The program was implemented in semi-urban municipalities in three regions of central Colombia, covering an area around the size of California. It had two components: psychosocial stimulation and micronutrient supplementation. The stimulation curriculum was based on the Jamaican home visiting model, which obtained positive short- and long-term effects (Grantham-McGregor et al. (1991), Walker et al. (2006, 2011) and Gertler et al. (2014)). The protocols designed by Grantham-McGregor et al. (1991) for Jamaica were adapted to be culturally appropriate for Colombia. The aims of the home visits were to improve the quality of maternal-child interactions and to assist mothers to participate in developmentally-appropriate learning activities, centered around daily routines and using household resources as learning tools.

We implement two key innovations vis-a-vis the Jamaican intervention with scalability and sustainability in mind. Indeed our program was specifically designed to go beyond the earlier small scale and tightly supervised efficacy trials. The first was that the intervention was implemented on a much larger scale than in Jamaica, covering a large part of the country and obtaining much larger sample sizes. The second was that the intervention was designed to be delivered by women drawn from the local community, with no specific prior professional experience.

To this end, home visitors were drawn from a network of local women, created by the administrative set-up of the CCT program. Every 50-60 beneficiaries elect a representative who is in charge of organizing social activities and acts as mediators

between them and the program administrators. These women, known as *Madre Líderes* (MLs), are beneficiaries of the program themselves. Given they are selected by their peers, one can deduce that they enjoy the trust of the community and are probably more entrepreneurial and proactive than the average beneficiary. In terms of specific characteristics they are on average about 10 years older (37) and have about one more year of education (8.5) than the subject mothers. Their score on a vocabulatory test is slightly higher than that of the mothers.[6] Finally, as mentioned in the introduction, another distinct feature of our intervention is that we targeted a more general poor population, namely the beneficiaries of the CCT program, which in Colombia is offered to the 20% poorest segment of the population, as compared to the extreme disadvantage of the malnourished population targeted by the Jamaican experiment.

The intervention was evaluated through a cluster randomized controlled trial involving the random allocation of 96 municipalities across central Colombia. After first stratifying into three large regions, 32 municipalities in each were randomly assigned to one of 4 groups: (i) psychosocial stimulation, (ii) micronutrient supplementation, (iii) both, and (iv) control. In each municipality, 3 MLs were selected and the children aged 12-24 months of the beneficiary households represented by each of these MLs were recruited to the study. There was a total of 1,429 children living in 96 towns in central Colombia. Possibly because the MLs are such trusted figures in their communities, compliance was high and the average number of home visits made was 63, which is 81% of those scheduled. The attrition rate between baseline and follow-up was around 10% across treatment arms, and the difference in loss among the groups was not statistically significant.[7]

As reported in Attanasio et al. (2014), there was no significant impact of micronutri-

[6]To measure vocabulary, we use the Peabody Picture Vocabulary Test (PPVT). The scores of the MLs was 28.2 versus 26.9 for the subject mothers in our sample. The difference has a p-value of 0.061.

[7]As we explain in Section 3.1, our data at baseline and at follow-up come from a household survey and from direct assessments administered to children in a community centre. The attrition rate for the household survey was 6.9%. The attrition rate for the direct assessments was 10.7%.

ent supplementation on any child developmental outcomes. In this paper, therefore, we focus on the psychosocial stimulation arm of the program and we refer to the "treated" group as those children who received the stimulation component of the intervention (groups i and iii) and to the "control" group as those children who did not (groups ii and iv).

Individuals randomized into our intervention were all eligible for and receiving subsidies from the CCT program. On average, households had been part of the CCT program for 21 months at baseline. This feature is common between treatment and control communities, but it is true that the context in which our program was implemented and in particular the existence of the CCT may be a factor in how effective the program was. This, of course, is related to the more general issue of extrapolating the effects of the program to other contexts outside the support of the data. Nevertheless, CCT programs are quite common in low-and-middle income countries and consequently the context is directly relevant to many other countries besides Colombia.

Finally, a frequently asked question is whether the intervention is just "teaching to the test" without leading to genuine advances in cognition. First, implementation of the curriculum has been shown to have long-run effects on cognition (Walker et al., 2005, 2011) and labour market outcomes (Gertler et al., 2014). This in itself is evidence that it can induce deep changes in achievement rather than just teach children to remember a few activities and perform better on a test. More generally, the intervention curriculum emphasizes cognitive, language and socio-emotional development through play and the promotion of mother-child interactions. While some of the play activities specifically address the type of cognitive and fine motor skills (building towers with blocks, tracing lines) and concepts (shapes, sizes, colors) that are assessed in developmental tests, the focus is on learning through play in a supportive and stimulating environment. Activities are introduced progressively and in developmental order to facilitate scaffolding - i.e. increasing or decreasing the challenge based on the child's performance - and

there is a strong emphasis on praising attempts and not only successes. The approach is aimed at promoting attention to task, perseverance and self-esteem; similarly, there is a strong focus on labelling the environment and looking at picture books together, which are activities that enrich vocabulary and promote bonding, attention (i.e. following a story) and other cognitive abilities (i.e. linking concepts, understanding cause and effect relations). All of these skills are associated with improved school readiness, school attainment and other outcomes linked to socio-economic success in life.

3 Data and measurement system

In this section, we describe the data we use, which was collected around the evaluation of the parenting intervention mentioned above. We then discuss an effective way of extracting the relevant information from such rich data with a measurement system that explicitly takes into account the relationship between relevant factors and available measures and the presence of measurement error.

3.1 Data

The main data we use in this paper comes from two rounds of data collection: before the intervention started (baseline) and just after it ended 18 months later (follow-up). In each round, information was collected in two ways: via a household survey in the home and via tests directly administered to children in a community centre. At the end of the paper, we also briefly discuss results from a second follow up (FU2), two years after the end of the intervention, although we do not use those data here.

The household surveys contain information on an extensive set of socio-economic and demographic characteristics, alongside a wealth of information around parenting, parental characteristics and maternal skills, including mothers' years of education, verbal ability, IQ, depressive symptoms and knowledge of child development. Among

others, among mothers, we administered the Raven's progressive matrices to test for IQ and the CES-D 10-item scale to assess depressive symptoms.

To measure children's developmental outcomes, we collected data based both on maternal reports and on direct assessment of the child. The measures of child development that we collected in the home setting via maternal report include: language development (that is, the number of words and complex sentences the child can say) using the vocabulary checklists in the Spanish Short-Forms of the MacArthur-Bates Communicative Development Inventories I and II (MacArthur); child temperament using Bates' Infant Characteristics Questionnaire (ICQ); and the attentional focusing and inhibitory control scales of the short versions of the Early Children's Behavior Questionnaire (ECBQ). All of these were measured using age-appropriate items pre- and post-intervention, with the exception of the ECBQ which was administered at follow-up only. In addition to these assessments via maternal reports, trained psychologists administered the Bayley Scales of Infant and Toddler Development III (Bayley) in community centres.[8] These direct assessments of the child took place over an average period of 1.5 hours and were aimed at measuring children's cognitive, language and motor development in depth.

The household surveys also contain detailed information on parental investments. We used a slightly modified version of the UNICEF Family Care Indicators (FCI) (Frongillo, Sywulka, and Kariger, 2003) which is based both on interviewer observations and maternal reports of the home environment. Specifically, this instrument includes interviewer observations of the types and numbers of play materials around the home and maternal reports of the types and frequency of play activities performed by the primary caregiver or any other adult older than 15 with the child in the last 3 days.

Examples of play materials include toys designed for learning shapes, toys that

[8]See Jackson-Maldonado et al. (2012) for MacArthur-Bates scales, Bates et al. (1979) for the ICQ and Putnam et al. (2006) for ECBQ and Bayley (2006).

induce physical movement, coloring books, and picture books. Examples of play activities include reading or looking at picture books together, telling stories, and labelling items in the home. Importantly therefore, the instrument affords the possibility of distinguishing between parental investments in 'materials' and in 'quality time', which are likely to have different costs and perhaps different (but possibly complementary) effects on child development.

As we want to assess the role of parental investment and distinguish it from the direct role that the intervention might play on child development, in measuring materials we instructed the data collectors to gather separate information about those materials (such as certain books and toys) that were directly provided by the intervention. When estimating the distribution of the factors measuring the two different types of parental investments, we use the measures that are not linked directly to the intervention.

Finally, in addition to the survey data collected around the evaluation of the intervention, we also use additional data sources to obtain information on municipality-level variables that we use as instruments. In particular, we use data on prices (of toys and food) and on maternal childhood exposure to violence. Appendix B provides details on all the measures of child development, maternal skills and parental investments collected as part of the survey and describes the auxiliary data sources we use to construct our instruments.

Appendix Table A.1 reports the baseline characteristics of children, their mothers and their households. At baseline, the children are on average aged 18 months. About 10% of them were born premature and 14% of them were stunted. On average, their mothers are 26 years old, have about 7.5 years of education and two-thirds of them are either married or cohabiting. There were no compromises to the randomization protocol and hence there is no reason to believe there is any bias. Most baseline characteristics are very well balanced including the baseline skills of the children. Although the mean of a few characteristics is significantly different between treated and controls

when tested individually (specifically among CESD scale items), none of these differences are significant at all when we allow for multiple hypothesis testing using the Romano and Wolf (2005) procedure.

3.2 Factor Models and the measurement system

Our main aim is to interpret the experimental results within the context of a model of parental investments and human capital production functions. To fix ideas, suppose we wish to estimate a production function for child skills:

$$\theta_{t+1} = f_{t+1}(\theta_t, I_{t+1}, P_t, X_t, \eta_{t+1}) \tag{1}$$

where θ_t and θ_{t+1} are vectors of the child's skills at t and $t+1$ respectively, I_{t+1} are parental investments that occur between the realizations of θ_t and θ_{t+1}, P_t are maternal skills measured at baseline, X_t is a vector of baseline household characteristics, such as household composition and η_{t+1} are random shocks to child development. The production function allows us to understand the pathways through which the experiment might affect outcomes: changes in parental investments and/or changes in the production function $f_{t+1}(\cdot)$, reflecting, for example, better use of parental inputs.[9]

As Cunha and Heckman (2008) explain, an important obstacle to estimating such a function is that the skills and investments are inherently unobservable. The various measures described in Section 3.1 can be viewed as error ridden indicators for these underlying latent factors. Using any one set of these measures in place of the latent factors could lead to severely biased results, whether the model is linear or not. We

[9]We use maternal skills as measured at baseline. However, we find no evidence of a treatment impact on any measures of cognitive skills or socio-emotional skills of the mother (the main primary caregiver in most households in our sample). This is in line with psychological evidence indicating that cognition (as measured by IQ) is rank stable by the age of 10 (Almlund et al., 2011). While it is more plausible that the intervention could have changed maternal socio-emotional skills, we find no such evidence. Had these maternal measures changed they could have been an additional channel of impact.

thus follow the approach of Cunha and Heckman (2008) and Cunha et al. (2010)[10] and develop a measurement system linking the observed measures to latent factors and estimate the distribution of such factors.

Suppose we have $\mathcal{M}^{\theta}_{kt}$ measures of child's skill θ^k_t of type k (e.g. cognitive or socio-emotional skills) in period t. Moreover, we also have $\mathcal{M}^{P}_{k}$ measures of maternal skills P^k of type k. Finally, we have $\mathcal{M}^{I}_{\tau t}$ measures of parental investments I^{τ}_t of type τ (e.g. time or material investments) made between $t-1$ and t. We denote m^{θ}_{kjt} the j-th measure of child's skill of type k at t, m^{P}_{kj} the j-th measure of maternal skill of type k, and $m^{I}_{\tau jt}$ the j-th measure of parental investment of type τ at t. As we estimate a different joint distribution of latent factors for the control and treated groups, in what follows we index the measures and latent factors by the treatment subscript d, where $d=0$ refers to the control group (no home visits) and $d=1$ refers to the treatment group (some home visits).

As is common in the psychometric literature, we assume a dedicated measurement system, that is one in which each measure only proxies one factor (Gorsuch, 1983, 2003). Although it is not necessary for identification, we maintain this assumption because it makes the interpretation of the latent factors more transparent and we find clear support for such a system in the data (see Appendix C). Assuming each measure is additively separable in the (log) of the latent factor it proxies,[11] we write the following system of equations mapping the j-th measure observed at some date t to the k-th

[10] More broadly this approach relates to the identification and estimation of nonlinear models with classical measurement error (Schennach, 2004, 2007).

[11] The measurement equations are specified in terms of the log latent factors. This ensures that the factors themselves only take positive values as required by the model.

latent (unobserved) factor for that date:

$$m^{\theta}_{kjdt} = \mu^{\theta}_{kjt} + \alpha^{\theta}_{kjt} \ln \theta^{k}_{dt} + \epsilon^{\theta}_{kjt} \quad (2)$$

$$m^{P}_{kjd} = \mu^{P}_{kj} + \alpha^{P}_{kj} \ln P^{k}_{d} + \epsilon^{P}_{kj} \quad (3)$$

$$m^{I}_{\tau jdt} = \mu^{I}_{\tau jt} + \alpha^{I}_{\tau jt} \ln I^{\tau}_{dt} + \epsilon^{I}_{\tau jt} \quad (4)$$

where the terms μ^{θ}_{kjt}, μ^{P}_{kj} and $\mu^{I}_{\tau jt}$ are intercepts, the terms α^{θ}_{kjt}, α^{P}_{kj} and $\alpha^{I}_{\tau jt}$ are factor loadings, and the terms ϵ^{θ}_{kjt}, ϵ^{P}_{kj} and $\epsilon^{I}_{\tau jt}$ are mean zero measurement error terms which are assumed independent of the latent factors and of each other.[12]

An assumption we have made in writing the system above is that the measurement system is invariant between treated and controls. This implies that any differences in the distribution of observed measures between the control and treated groups result from differences in the distribution of the latent factors only. As we show in Appendix D, none of our results are sensitive to whether we allow for the measurement treatment to be affected by the treatment or not.

Because the latent factors are unobserved, identification requires normalizations to set their scale and location (Anderson and Rubin, 1956). We set the scale of the factors by setting the factor loading on one of the measures (say the first) of each latent factor to 1, that is: $\alpha^{\theta}_{k1t} = \alpha^{P}_{k1} = \alpha^{I}_{\tau 1t} = 1$, $\forall t, \tau = \{M, T\}$ and $k = \{C, S\}$. When it comes to the child's skills, we normalize the factor loading on the same measures at baseline and follow-up.[13] We set the location of all the factors by fixing the mean of the latent factors in logs to 0 for the control group; the difference between the treatment group's

[12]The assumption that the errors are independent of each other can be relaxed somewhat. Some of the child cognitive outcomes, for example, are based on child level observations and are collected by a trained psychologist in community centers, while others are based on maternal reports and are collected in the home (on a different day) by a different interviewer. However, it is certainly possible that measurement errors are correlated, even in this case from say child behavior, the implications of which should be studied in future research.

[13]For cognitive skills, we define the scale based on the Bayley cognitive score both at baseline and follow-up. For socio-emotional skills, we normalize the factor loadings on the item measuring difficulty in child's temperament in the ICQ.

location and that of the control group (which is set to zero) is taken to be the average effect of the treatment.

With the assumptions and normalizations already made and based on the Kotlarski theorem and further extensions, Cunha et al. (2010) show that both the distribution of measurement errors and the latent factor distribution are non-parametrically identified so long as we have at least three measures with nonzero factor loadings corresponding to each latent factor.[14] While these assumptions are sufficient for identification, some of them could be relaxed as shown in Cunha et al. (2010).[15] For instance, the same measure could be allowed to load on several factors, provided there are some dedicated measures. It would also be possible to allow measurement error to be correlated across measures of the same factor, as long as there is one measure whose error is independent from those of other measures of the same factor.

A question of practical importance relates to the scale of the latent factors and what they actually mean for measures of interest such as earnings. This is the issue of anchoring discussed in Cunha et al. (2010) who provide a theoretical treatment.[16] In our paper, we normalize the cognitive factor in both periods to the Bayley cognitive scale. This has a cardinal interpretation (the number of tasks completed correctly), and the same test is applied across different ages (up until 42 months), allowing for comparability. For socio-emotional skills we also normalize to the same ICQ item (whether the child is difficult) in both periods. The lack of long-term longitudinal data prevents us from converting these units to future earnings or other adult outcomes of interest.

[14] See also Schennach (2004), Schennach (2007), Hu and Schennach (2008), Carneiro, Hansen, and Heckman (2003), Heckman, Pinto, and Savelyev (2013) and Cunha and Heckman (2008).

[15] See also Carneiro et al. (2003) and Cunha and Heckman (2008)

[16] Cunha et al. (2010) provide a general theoretical treatment of anchoring and in their main empirical results they anchor the measure of skills measured at the oldest age to years of education. They then assume that the same anchoring scale applies to measures of cognition and socio-emotional skill measured at earlier ages. Nielsen (2015) discusses using ordinal tests scores to measure achievement gaps, and Agostinelli and Wiswall (2016) discuss how rescaling in a multiperiod production function can lead to biases in the estimation of the substitution elasticity.

Table 1: Measurement system

Latent factor	Measurement	Survey	% Signal	
			Controls	Treated
Child's cognitive skills at FU (θ^C_{t+1})	Bayley: cognitive	FU	78%	79%
	Bayley: receptive language	FU	75%	76%
	Bayley: expressive language	FU	78%	79%
	Bayley: fine motor	FU	59%	61%
	MacArthur: words the child can say	FU	64%	65%
	MacArthur: complex phrases the child can say	FU	52%	54%
Child's cognitive skills at BA (θ^C_t)	Bayley: cognitive	BA	70%	70%
	Bayley: receptive language	BA	73%	72%
	Bayley: expressive language	BA	75%	74%
	Bayley: fine motor	BA	60%	59%
	MacArthur: words the child can say	BA	45%	44%
Child's socio-emotional skills at FU (θ^S_{t+1})	ICQ: difficult (-)	FU	74%	71%
	ICQ: unsociable (-)	FU	33%	30%
	ICQ: unstoppable (-)	FU	59%	55%
	ECBQ: inhibitory control	FU	73%	69%
	ECBQ: attentional focusing	FU	27%	24%
Child's socio-emotional skills at BA (θ^S_t)	ICQ: difficult (-)	BA	68%	71%
	ICQ: unsociable (-)	BA	28%	31%
	ICQ: unadaptable (-)	BA	35%	38%
	ICQ: unstoppable (-)	BA	22%	25%
Material investment at FU (I^M_t)	FCI: no. of different types of play materials	FU	94%	97%
	FCI: no. of coloring and drawing books	FU	17%	29%
	FCI: no. of toys to learn movement	FU	61%	76%
	FCI: no. of toys to learn shapes	FU	69%	82%
	FCI: no. of shop-bought toys	FU	61%	76%
Time investment at FU (I^T_t)	FCI: no. of different types of play activities in last 3 days	FU	87%	93%
	FCI: no. of times told a story to child in last 3 days	FU	66%	81%
	FCI: no. of times read to child in last 3 days	FU	73%	85%
	FCI: no. of times played with toys in the last 3 days	FU	55%	72%
	FCI: no. of times named things to child in last 3 days	FU	56%	73%
Mother's cognitive skills at BA (P^C)	Mothers' years of education	FU	54%	50%
	Mother's Raven's score (IQ)	BA	54%	51%
	Mother's vocabulary	FU2	65%	62%
	FCI: no. of books for adults in the home	BA	39%	36%
	FCI: no. of magazines and newspapers in the home	BA	20%	19%
Mother's socio-emotional skills at BA (P^S)	CESD: did you feel depressed? (-)	BA	70%	73%
	CESD: are you bothered by what usually don't? (-)	BA	42%	45%
	CESD: did you have trouble keep mind on doing? (-)	BA	49%	52%
	CESD: did you feel everything you did was an effort? (-)	BA	45%	49%
	CESD: did you feel fearful? (-)	BA	47%	51%
	CESD: was your sleep restless? (-)	BA	34%	38%
	CESD: did you feel happy?	BA	46%	50%
	CESD: how often did you feel lonely last week? (-)	BA	51%	55%
	CESD: did you feel you couldn't get going? (-)	BA	48%	52%

Note: This table shows the measures allowed to load on each latent factor, as well as the fraction of the variance in each measure that is explained by the variance in signal, for the control and treatment groups separately. "BA" refers to Baseline, "FU" refers to the first-follow-up survey and "FU2" refers to the second follow-up survey collected 2 years after the intervention ended. The symbol (-) indicates that the scoring on these measures was reversed so that a higher score on the corresponding latent factor means a higher level of skill.

3.3 Specification of the measurement system

To implement the measurement system above, we first perform an exploratory factor analysis, reported in Appendix C, to identify in a preliminary step the relevant measures and their allocation to factors. We then allocate measures observed in the data to particular factors, as is shown in Table 1. The factor loading on the first measure is normalized to one and thus this measure defines the scale of the latent factor.

As reflected in the table, we did not necessarily use the same set of measures of the child's skill at baseline and at follow-up, the main reason being that we only included age-appropriate items that provide relevant information about the latent skill. For example, the MacArthur item measuring the number of complex phrases a child can say is too advanced for children at 1-2 years old and hence was only administered at follow-up when children were between 2.5 to 3.5 years old. Similarly, with respect to socio-emotional skills, the ECBQ is designed to measure temperament among children aged 3-7 and therefore was only administered at follow-up.[17] However, in both rounds, we use the same measure to normalize the child's baseline cognitive and socio-emotional skills.

In our model we use mother's skills to control for parental background. During the data collection process, we had to focus only on the mother's skills (who is almost always the principal caregiver and often a single mother) because of resource constraints and in order to keep interview times at a reasonable level. In so doing, it is possible that we miss the influence of the father; however, we expect to be capturing at least some of that by conditioning on the baseline skills of the child. We use baseline measures to extract two factors measuring the mother's cognitive and socio-emotional skills, with the exception of the vocabulary test, which was administered at follow up and the Raven's score which was administered at a later round of data collection (2 years after

[17]The ICQ is in principle designed for children up to 2 years old. We administered the same questions of the ICQ at baseline and follow-up after consultation with the developer of the test.

the end of the intervention). In both cases we checked and the intervention had no impact on the scores.

The parameters of the measurement system for treatment and control are estimated together with the latent factor distributions as described above. To do so, we use the estimation method described in Attanasio et al. (2017), which approximates the joint distribution of the latent factors by a mixture of normals (as in Cunha et al. (2010)) and the measurement error distribution by a normal distribution. We report estimates of the factor loadings and distribution of measurement errors in Appendix C. To assess the extent of information relative to measurement error contained in each of the measures, we compute the signal-to-noise ratio measuring the fraction of the variance of each measure driven by signal. For example, for the j-th measure of child's skills of type k, this ratio is defined as:

$$s_j^{\ln \theta^k} = \frac{(\alpha_j^k)^2 \; Var(\ln \theta^k)}{(\alpha_j^k)^2 \; Var(\ln \theta^k) + Var(\epsilon_j^k)}$$

where we have assumed that the j-th measure of latent factor θ^k can be written, simplifying notation, as:

$$m_j^\theta = \mu_j^k + \alpha_j^k \ln \theta^k + \epsilon_j^k$$

The last two columns of Table 1 report the signal-to-noise ratio for each of the measures used in the analysis for the control and treated groups separately. These numbers can be different because the joint distribution of latent factors is allowed to be different between the two groups. Clearly, there is much variation in the amount of information contained in each measure of the same factor. For example, 78% of the variance in the *Bayley: Cognitive* item is due to signal, whereas only 51% of the variance in the *Mac Arthur: Complex Phrases* item is due to signal. Overall, most measures are far from having 100% of their variance accounted for by signal,

which illustrates the usefulness of the latent factor approach in modeling human capital accumulation and parental investments: without such an approach, one would risk to obtain severely attenuated coefficients, masking the importance of investments and background variables on child development.

4 Short-term impacts on child outcomes and parental investments

In this section, we document the impacts of the intervention on child's cognitive and socio-emotional development as well as parental investments, observed at first follow-up, just after the 18 month-long intervention ended. The latter provide the basic input to perform the mediation analysis discussed below to uncover the mechanisms behind the observed impacts on children outcomes. We focus on the impact of the psychosocial stimulation component of our intervention because there were no significant impact of micronutrient supplementation on any child developmental outcomes (Attanasio et al., 2014).[18]

4.1 Impacts on child development

Each panel of Table 2 reports the estimated impacts of receiving the home visits on one of four sets of outcomes: (i) cognitive development; (ii) socio-emotional development; (iii) parental investment in play materials; (iv) parental investment in play activities. In addition to the impact on each measure, we also report the impact on the mean of the corresponding log latent factor. The results in the first panel imply an increase of 0.25 of a standard deviation (SD) in cognitive development and an increase of 0.175 SD

[18] If we explicitly control for the fact that half of the stimulation group also received micronutrient supplementation, the impact on cognition and receptive language remains virtually the same, with a very small increase in the point estimates we report below (see Appendix Table D.1).

Table 2: Treatment impacts on raw measures and factors

	Treatment effect		
	Point estimate	Standard error	Sample size
A - Child's cognitive skills at follow-up			
Bayley: cognitive	0.250	(0.063)	1,264
Bayley: receptive language	0.175	(0.063)	1,264
Bayley: expressive language	0.032	(0.062)	1,263
Bayley: fine motor	0.072	(0.060)	1,262
MacArthur: words the child can say	0.092	(0.064)	1,322
MacArthur: complex phrases the child can say	0.058	(0.054)	1,322
Cognitive factor	0.115	(0.051)	
B - Child's socio-emotional skills at follow-up			
ICQ: difficult (-)	-0.041	(0.054)	1,326
ICQ: unsociable (-)	-0.075	(0.045)	1,326
ICQ: unstoppable (-)	-0.032	(0.054)	1,326
ECBQ: inhibitory control	-0.003	(0.058)	1,323
ECBQ: attentional focusing	0.070	(0.049)	1,323
Socio-emotional factor	0.087	(0.044)	
C - Material investment at follow-up			
FCI: no. of different types of play materials	0.215	(0.064)	1,326
FCI: no. of coloring and drawing books	-0.133	(0.056)	1,326
FCI: no. of toys to learn movement	-0.048	(0.065)	1,326
FCI: no. of toys to learn shapes	0.416	(0.088)	1,326
FCI: no. of shop-bought toys	0.024	(0.061)	1,326
Material investment factor	0.227	(0.069)	
D - Time investment at follow-up			
FCI: no. of different types of play activities in last 3 days	0.277	(0.050)	1,326
FCI: no. of times told a story to child in last 3 days	0.138	(0.060)	1,326
FCI: no. of times read to child in last 3 days	0.362	(0.062)	1,326
FCI: no. of times played with toys in last 3 days	0.175	(0.056)	1,326
FCI: no. of times named things to child in last 3 days	0.137	(0.048)	1,326
Time investment factor	0.302	(0.068)	

Note: All scores have been internally standardized non-parametrically for age and are expressed in standard deviation units (see Appendix B for details about the measures and the standardization procedure). The effects relating to the *latent factors* are in log points. Coefficients and standard errors clustered at the municipality level (in parentheses) from a regression of the dependent variable measured at follow-up on an indicator for whether the child received any psychosocial stimulation and controlling for child's sex; tester effects and baseline level of the outcome.

in receptive language, assessed using the Bayley.[19] The cognitive factor summarizing

[19]These treatment effects are slightly different from those reported in Attanasio et al. (2014) because in this paper we estimate the impact of psychosocial stimulation by pooling the two groups that

all these effects shows a substantial and significant increase of 11% (0.11 log points) amongst the treated group relative to the control group. The second panel of the table also shows that the intervention led to an overall improvement in socio-emotional development (p-value<0.05).

In Figure 1, we plot the estimated densities of some of the factors for the control and treated groups and perform a Kolmogorov-Smirnov (K-S) test of the hypothesis that the corresponding CDFs are equal to each other (the p-values of the tests are reported in the figure and have been derived using the bootstrap).[20] The first two panels show the distribution, in treatment and control villages, of cognitive and socio-emotional skills at baseline. The two densities overlap each other and the K-S test cannot reject that they are equal to each other, thus confirming that our sample is balanced. The following two panels depict the distribution of cognitive and socio-emotional factors at follow-up. In the case of the cognitive factor, we see that the shift in the mean reported in Table 2 reflects a shift in the entire distribution. For the socio-emotional factor, however, the shift occurs mainly for children below the median.

4.2 Suggestive evidence on the mechanisms behind the impacts

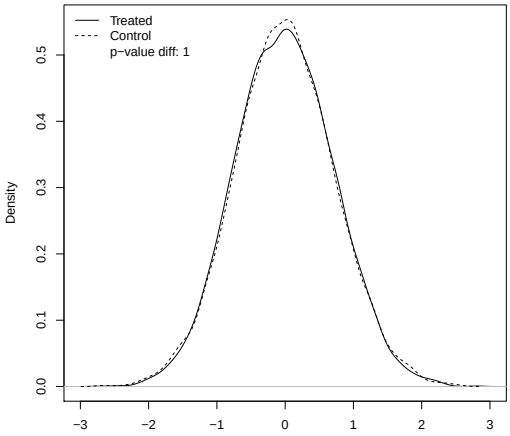
In the last two panels of Figure 1, we notice a strong shift to the right of the distributions of both the material and time investment factors. For either type of investments, the K-S strongly rejects the equality of the corresponding densities between control and treated groups. The bottom two panels of Table 2 focus on the mean impacts of the stimulation intervention on parental investments and indicate substantial impacts on several individual items, as measured by the Family Care Indicator (FCI), as well as on the two latent factors measuring investments. Panel D of Table 2 shows that all types of time activities increase, but among play materials the increase is not uniform

received it and the two groups that did not, while Attanasio et al. (2014) estimates the impact of each of the four arms of the intervention separately.

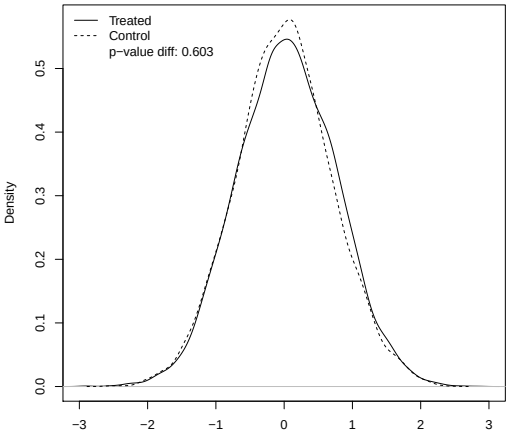
[20]The estimation method used for this purpose is based on Attanasio et al. (2017)

Figure 1: Kernel densities of latent factors

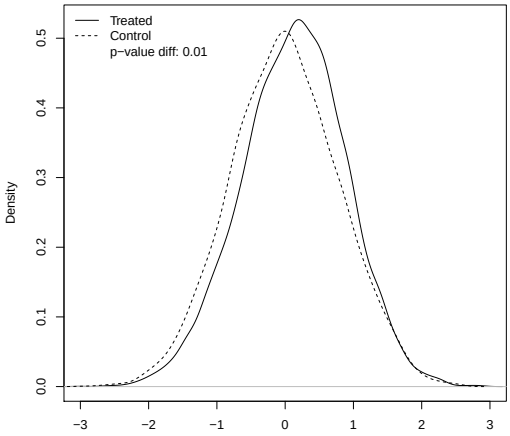
(a) Children's cognitive skills, baseline



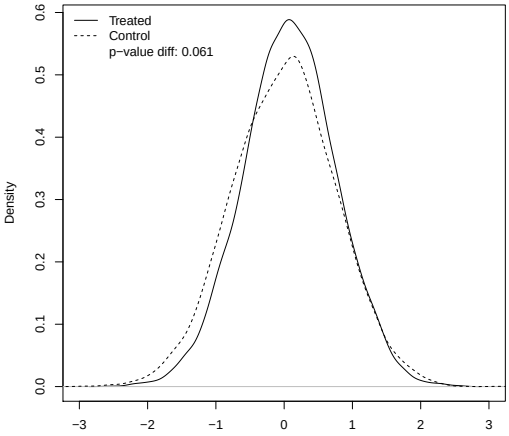
(b) Children's socio-emotional skills, baseline



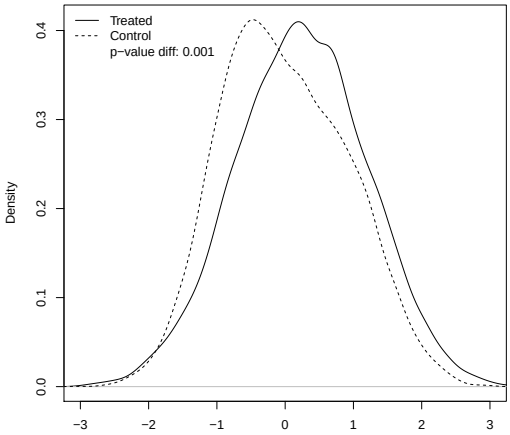
(c) Children's cognitive skills, follow-up



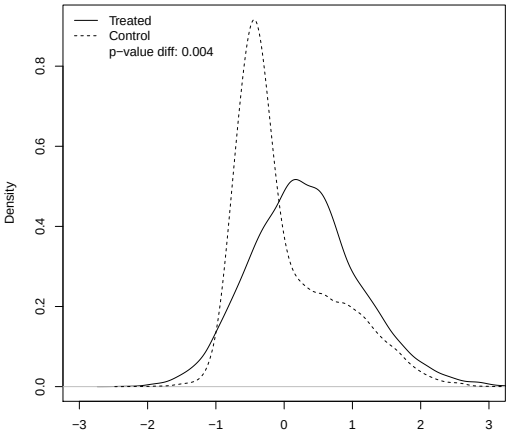
(d) Children's socio-emotional skills, follow-up



(e) Material investments - follow-up



(f) Time investments - follow-up



Note: These kernel densities are constructed using 10,000 draws from the estimated joint distribution of latent factors for the control group and for the treated group. For each factor, we perform a Kolmogorov-Smirnov test using the bootstrap and accounting for the entire estimation procedure. p-values reported in each panel.

(Panel C). Specifically, there is an increase in most toys but a reduction in coloring books, which may reflect crowding out specially because the home visitors intentionally left picture books behind. The overall material investment factor registers a highly significant increase however.

As mentioned above, the measure of materials relates specifically to items provided by the parents. Although the home visitors were supposed to take away all intervention toys (with the exception of picture books) upon the completion of the last home visit, some were left behind at the end of the intervention as is evident from the summary statistics reported in Appendix Table D.3.[21] However, since we are able to separately measure parental contributions from intervention materials, this does not pose any problem for our analysis, and whenever we refer to material investments we exclude intervention play materials and only keep items provided by the parents. The impact of the intervention materials, if any, will be captured by the treatment dummy; we return to this point when interpreting our results.

As we show above, the measures of parental inputs relating to materials and quality time both increased. As argued by DelBoca, Flinn, and Wiswall (2014) it may be important distinguish these inputs because they can have different impacts on child development. On the other hand the inputs are clearly related and the two factors are correlated (0.64). Thus one could imagine a more parsimonious approach where they are combined into one investment factor. However, another good reason for keeping them separate relates to the way they are measured. Materials are actually observed and enumerated by the surveyors; quality time items are measured by maternal self-reports, which may make them noisier measures of parental investment. This may also make them more likely to be subject to intervention bias due to the fact that mothers in

[21]The median number of days between the end of the intervention and data collection was 10. Almost all interviews were completed within 40 days. Very few households were interviewed a few days before the end of the intervention. Omitting these households from the analysis leaves the results completely unchanged.

the treatment group may exaggerate the extent to which they engage in developmental activities with the children. Without this implying that there is no information in the time measure it may be a less reliable measure of parental behavior than material investments, which are directly observed and also likely to reflect actual effort by the parents. We will return to this important issue when interpreting the results.

5 The accumulation of human capital in the early years and the role of the intervention

To better understand the determinants of early childhood development and to explore the way that the intervention affected outcomes we now specify a model of parental investments and child skill formation, where skills take two dimensions, namely cognitive and socio-emotional skills. We use such a model to inform the mediation analysis aimed at explaining the channels through which the intervention generated the impacts documented in Section 4. An important element of the model is that parents can choose to invest in materials and quality time.

We refer to the baseline period as t, when children were between 12 to 24 months old, and to the post-intervention period as $t+1$, when children were between 30 to 42 months old. Child skills at $t+1$ are assumed to be a function of the vector of child skills at t, maternal skills at t, parental investments in the intervening period and random shocks. However, rather than modeling investment choices resulting from the dynamic optimization of a household problem as in DelBoca, Flinn, and Wiswall (2014), we estimate a pair of reduced form investment equations, which can be interpreted as an approximation to those derived (numerically) in a full structural model. By not imposing all the restrictions from a specific structural model we do not have to take a stance on whether parents know the process of child development reflected in the structure of the production function.

The model we use allows us to characterize the process of early child development and provides a framework to understand the mechanisms that generated the intervention's impacts. One mechanism through which the intervention may have operated is by changing the production function itself. On the one hand, the stimulation provided during the home visits may be a new input in the development of the child, and this would be captured by a shift in Total Factor Productivity (TFP) parameter or other parameters of the production function. On the other hand, parents, now guided by the intervention, may use time and resources in a more effective way. This interpretation implies that, despite the richness of our data, some aspects of investment quality may not be captured by our measures and thus get embodied in the estimates of the production function parameters.[22]

A second mechanism through which the intervention could generate impacts on child development could be an increase in parental investments. Indeed, the intervention aims to strengthen child-mother interactions and encourage mothers to engage more with the child by incorporating age-appropriate play activities in the daily routine, introducing new toys constructed with home-made materials and spending time reading, telling stories or singing. However, it is also possible that investments could decline as parents shift their attention and resources elsewhere (for example, to other children) because they perceive the intervention itself as some form of investment either in itself (effectively an in-kind transfer). Such crowding-out of private resources is a standard concern in programs that target children.

Finally, the intervention could also have affected maternal cognitive or, more plausibly, socio-emotional skills. Many of the mothers (37%) were depressed at baseline

[22]We made every effort to collect both time and resource use carefully targeted to the child with an emphasis on items that can drive development. For example, one of our measures is the number of times spent reading with the child in the last 3 days. Yet, it is still a possibility that as a result of the intervention, parents may be more able to select age-appropriate or stimulating stories to read with their child. Our measure of the frequency with which parents read with their child would not pick up this change in the quality of interaction, which would instead be picked up by a shift in the productivity of time investments.

according to the CESD scale, and it is plausible that the treatment mitigated this. Although we checked for such impacts, we did not detect any differences in our measures of maternal skills (either cognitive or socio-emotional skills) between the control and treatment groups after the intervention; thus this potential change is not a mechanism that contributed to the outcome. In our estimated model we only include baseline maternal skills.[23]

5.1 The production function for human capital

We consider a production function of human capital that maps initial conditions, parental investment of different types and other factors on two different dimensions of child development. In particular, we assume that the stock of skills of child i in period $t+1$ is determined by the vector of child's baseline cognitive and socio-emotional skills θ_{it} embodying the initial conditions at the time of observation (possibly including any paternal influence), the mother's cognitive and socio-emotional skills denoted by P_{it}^C and P_{it}^S respectively, and the investments I_{it+1} made by the parents between t and $t+1$. We also allow for the effect of a variable η_{it+1}^k that reflects unobserved shocks. As with skills, parental investments I_{it+1} can be a multi-dimensional vector. We denote material investments by I_{it+1}^M and time investments by I_{it+1}^T.

Following our own earlier experimentation, we assume the production function for each of the two skills is Cobb-Douglas, so we can write the technology of formation for skill k as follows:[24]

$$\begin{aligned} ln(\theta_{idt+1}^k) =& A_d^k + \gamma_{1d}^k ln(\theta_{it}^C) + \gamma_{2d}^k ln(\theta_{it}^S) + \gamma_{3d}^k ln(P_{it}^C)^+ \gamma_{4d}^k ln(P_{it}^S) \\ &+ \gamma_{5d}^k ln(I_{idt+1}^M) + \gamma_{6d}^k ln(I_{idt+1}^T) + \gamma_{7d}^k n_{it} + \eta_{it+1}^k \qquad k \in \{C, S\} \end{aligned} \tag{5}$$

[23] The effect of the intervention on the principal component factor of the CES-D scale items at follow-up is 0.13 of a standard deviation (with a p-value of .12), which, given the way the factor is defined, is indicative of an improvement but too insignificant to rely upon.

[24] Cunha et al. (2010) use a CES, while Cunha and Heckman (2008) use a log linear specification.

where C and S stand for cognitive and socio-emotional skills respectively. The term n_{it} is the number of children in the household and allows for the possibility that the presence of siblings affects child development because of spillover effects and more broadly because of the learning and socialization that can be achieved by interacting with other older children.[25] It is possible, on the other hand, that the presence of siblings dilute attention and resources. As we discuss below, such an effect could be captured by the investment functions. A_d^k is a factor-neutral productivity parameter or TFP and depends on the treatment status of the child (d) to capture the potential direct effect of the home-visitor stimulation during her weekly visit. Finally, all the parameters are specific to a particular skill.

5.2 Parental investments

We model investments as a function of the child and the mother's baseline skills and the number of children in the household.[26] The number of children in the household may dilute both resources and time devoted to the subject child. We also include a vector of variables Z_{it}, which determine investments but do not enter the production function. We discuss them below. The investment equations we estimate are

$$\begin{aligned}\ln(I^{\tau}_{idt+1}) =& \lambda^{\tau}_{0d} + \lambda^{\tau}_{1d}\ln(\theta^{C}_{it}) + \lambda^{\tau}_{2d}\ln(\theta^{N}_{it}) + \lambda^{\tau}_{3d}\ln(P^{C}_{i}) + \lambda^{\tau}_{4d}\ln(P^{S}_{i}) \\ &+ \lambda^{\tau}_{5d}\ln(n_{it}) + \lambda^{\tau}_{6d}Z_{it} + u^{\tau}_{it+1}, \qquad \tau = \{M, T\}\end{aligned} \tag{6}$$

As implied by the subscript d all coefficients could change with the treatment, a hypothesis we directly test. The effect of background variables on parental investment, given child initial conditions, is an important potential source of socio-economic gradients in child development. Moreover, the extent to which investments increase with child initial abilities is a reflection of parental beliefs about the heterogeneity of returns

[25] Since our subject children are 12-24 months old at baseline these are almost always older children.
[26] The measure includes the subject child so the minimum is 1.

to such investments as well as parental taste for redistribution among children.

5.3 Estimation and mediation analysis

Parental investments are an input in the production function. However, they may be endogenous, i.e. it may be that $E(\eta^k_{it+1}|I^\tau_{it+1}) \neq 0$. In particular, parental investments might respond to unobserved, time-varying shocks in order to compensate or reinforce their effects on child development. Consider, for example, the case of a child who is suddenly affected by a negative shock, such as an illness, which is unobserved to the econometrician but perceived by the parents as delaying the child's development. As a result of this shock, parents might decide to invest in their child's development more than they would have otherwise. This parental response would create a negative correlation between parental investments and the unobserved error η^k_{it+1}, biasing downwards the impact of investments. Alternative assumptions about preferences and technologies (or technologies as perceived by the parents) can create different patterns of correlations between shocks and investment and, therefore, introduce different types of biases.

Standard mediation analysis, as in Heckman et al. (2013), takes all inputs as exogenous. It then considers various possible channels through which the intervention could affect outcomes and tries to establish which of them can explain the observed impacts on the outcomes. In our case, this approach would correspond to estimating the production functions by OLS, allowing the intervention to affect outputs (cognition and socio-emotional skills) directly, as well as indirectly through its impact on investments. One could then decompose the overall effect of the intervention into a direct effect, which could be interpreted as an improvement in productivity, and an indirect effect mediated by the increase in investment. Such an approach, however, can lead to misleading conclusions if investments are endogenous: if, for instance, ignoring endogeneity leads to under-estimate the impact of investment on outputs, then the channels

through which the intervention can affect outcomes will be potentially misinterpreted. To deal with the endogeneity of investment we use instrumental variables. We discuss our choice of instruments in the next subsection.

To estimate the model we proceed in three steps. First, as mentioned earlier, we estimate the covariance structure and the factor loadings of the latent factors based on the covariance structure of the observed measures. This step requires no distributional assumptions on the latent factors or the measurement error, but only relies on the restrictions embedded in the measurement system and discussed earlier. For estimation however, we assume that the latent factors are distributed as a mixture of two normal distributions and that measurement error is normally distributed. In the second step, we use the estimates of the measurement system to predict Bartlett factor scores for each individual in the data. In the final and third step, we use these predicted scores as observable data to estimate the parental investment and production function equations.

The third step requires correcting for the measurement error introduced from the fact that we use predicted values of the latent factors instead of the actual ones. To do so, we adapt the correction method described in Heckman et al. (2013) in the context of instrumental variables. Using the predicted factor scores and instruments, we estimate the investment equations and the reduced form of the production functions, where we obtained the latter by substituting material and time investments in the production function with their relevant first stage equations. We correct the investment and reduced form coefficients using the method described in Heckman et al. (2013) and finally recover the structural coefficients of the production function using a minimum distance estimator (Rothenberg, 1971).

We compute 95% confidence intervals and critical values for test statistics using the cluster bootstrap, where the entire estimation procedure is replicated 1000 times. We cluster at the municipality level which was the randomization unit in our experiment.

5.4 Choice of instruments

When investments are treated as endogenous, identification requires instruments that are relevant and can be excluded from the production function, under reasonable assumptions. A potential instrument is the intervention itself, which was randomized. However, the fact that we wish to test whether the intervention affected the production function directly as one of the possible channels through which it operated precludes us from using it as an identifying variable that is excluded from the production function. Moreover, because we have two endogenous variables (material and time investments), we need at least two instruments so the randomization alone would not be enough to identify the model anyway.

Consistent with a standard model of parental investment, we expect material investments to be related to the prices of relevant goods. Specifically, we use the average price of toys and the average price of food items in the municipality of residence. We assume that the variability of prices across communities is unrelated to factors affecting the development of cognition and socio-emotional skills of children.

Finding instruments for *quality time* is more challenging. This input may reflect more the way parents spend time with their children and the type of activities they engage in than the amount of time spent. This intuition is confirmed by results in Table 3, which show that the intervention had no impact on maternal labor supply: the impact of the treatment on both employment and weekly hours is small and insignificant, although it had large effects on time investments. This evidence suggests that the margin of adjustment is not between work and quality time with children, but rather between quality time versus other household production activities or leisure, which are excluded from our measures of time investments. Consequently, measures of the opportunity cost of time, such as village-level female or male wages, are unlikely

to have much explanatory power.[27] Below we also show that quality time investment is also unrelated to the relative prices we use as instruments for material investment.

Table 3: Impact of the intervention on maternal labor supply

	Employment	Hours of work
Treatment dummy	-0.0373	0.208
	(0.0852)	(1.793)
Demographic controls		
Mother's age	0.193***	3.839***
	(0.0439)	(0.802)
Mother's age squared	-0.003	-0.0542***
	(0.0007)	(0.0130)
Years of education	0.056***	0.921***
	(0.0121)	(0.214)
No. of children = 2	-0.213**	-4.429**
	(0.0983)	(1.866)
No. of children = 3	-0.397***	-5.905**
	(0.124)	(2.382)
No. of children = 4	-0.230	-2.057
	(0.166)	(3.158)
No. of children = 5	-0.664***	-13.71***
	(0.236)	(3.612)
No. of children > 6	-0.346	-0.169
	(0.244)	(5.107)
Constant	-3.457***	-48.39***
	(0.659)	(11.91)
Observations	1,210	1,200

Note: Interviewer fixed effects included. Default number of children is one, which is the subject child. Asymptotic standard errors in parentheses allowing for clustering at the municipality level (randomization unit).

To instrument time investments, we instead focus on variables that have the potential to affect the willingness and ability of mothers to engage with their children. Specifically, we exploit the fact that Colombia has a long and well documented history of civil conflict that has affected large parts of the country and, in particular, rural areas. It is well documented that exposure to violence can cause emotional de-

[27] In earlier versions we demonstrate that indeed they do not.

tachment, which can impede or make subsequent interaction with one's own children harder (Betancourt, 2015; Creech and Misca, 2017, for related evidence). This leads us to consider maternal exposure to past conflict as a potential instrument for current quality time investment, where we exploit variation in the intensity of conflict across municipalities.[28]

As with our price instruments, our identifying assumption is that the variation in conflict across municipalities is orthogonal to unobservable factors affecting child development. Although this assumption is not testable, in Table 4 we provide empirical evidence that is strongly suggestive of its validity: we show that the incidence of conflict is not associated with mothers' baseline characteristics or with household wealth. In the first column, we use measurements of variables related to maternal skills as explanatory variables, while in the second column we use factor scores for maternal skills, which summarize more information and improve the power of our test.

While the conflict itself was mainly over in the sampled communities by the time of our experiment, there were sporadic terrorist attacks during the period of the intervention. To check that our measure of maternal childhood exposure to conflict is not related to current violence (which could have a direct impact on the child), we also include a measure of such attacks in these regressions. The results in Table 4 strongly suggest that they are not. Thus, as is evident from these results, the past incidence of conflict is not associated with baseline characteristics relevant for child development. Individually all variables are insignificant, and the p-values for the joint significance of the right hand side variables are 0.46 and 0.35 in each regression respectively. However, as we show below, mothers' childhood exposure to conflict is a strong determinant of the quality time that they spend with their child.

[28]Specifically, our instrument is defined as the number of conflicts against the civil population divided by population (in thousands) in a given municipality when the mother was a child.

Table 4: Conflict and characteristics of mothers and children at baseline

	Raw Data	Factor Scores
Mother's years of education	-0.001	
	(0.001)	
Mother's CESD - "Did you feel depressed?" (-)	0.004	
	(0.003)	
Mother's cognitive factor		-0.005
		(0.003)
Mother's socio-emotional factor		0.006
		(0.005)
Household's wealth index	0.003	0.002
	(0.004)	(0.004)
Terrorist attacks between conception and baseline	0.004	0.005
	(0.008)	(0.008)
Constant	0.051	0.051
	(0.015)	(0.014)
Observations	1,108	1,023
R-squared	0.007	0.009
Significance of all explanatory variables:		
F-stat	0.912	1.125
p-value	0.461	0.350

Note: Dependent variable is the number of conflicts against civil population/1000 population when the mother herself was a child. All explanatory variables are measured at baseline. Asymptotic standard errors in parentheses allowing for clustering at the municipality level (randomization unit).

6 Results

We now report our empirical results. We start with estimates of the investment function before moving on to estimates of the production functions. The latter allow us to investigate what generates the impacts of the intervention on child development, a key question that impinges on the design of such programs.

6.1 Estimates of the investment functions

The estimates of the investment equations are reported in Table 5. The first two columns of the table report estimates of the material and time investment equations, where we use toy price, food price and maternal childhood exposure to conflict as exclusion restrictions. The third column reports estimates of the material investment equation, where we only include prices as exclusion restrictions. This first stage will be used to estimate a production function for cognitive skills that does not include time investments as an input (and that will turn out to be our preferred specification).

The first striking result is the impact of treatment on investments: it increases materials by 20% and time by 33%, and both effects are highly significant. The results reported in Table 5 exclude interactions of the treatment parameter with the remaining variables, which were found to be insignificant.[29] These estimates of the impact of the intervention on investments are driven by the experimental design and do not require any of the assumptions necessary for the identification of the production functions.

The fact that the intervention increased the quality time and resources that parents provide to children is important because it shows that parents are willing to reinforce the intervention. While we already showed some evidence of crowding out for individual items, overall the opposite seems to be happening. From a policy perspective this is a major conclusion that should encourage further interventions in such contexts. As we shall see below, this increase in parental investment is the key source of success of the intervention.

Turning now to the other regressors, we find that parents invest more resources in children with a higher baseline level of cognition (elasticity of 0.13) but the child's baseline socio-emotional skills have no impact on either type of investment. The elasticity of both material and time investments with respect to maternal cognition is very high

[29]The estimates where all parameters of the investment functions are allowed to vary with treatment are shown in Web Appendix Table E.1. We test the joint significance of the interaction terms and find that we cannot reject that all the interactions are equal to 0 for both material and time investments: the p-value for the material investment equation is 0.369 and the p-value for the time investment equation is 0.099.

Table 5: Estimates of the material and time investment equations

	Instruments: Prices and Conflict		Instruments: Prices only
	Material investment	Time investment	Material investment
Intercept	-0.016	0.001	-0.016
	[-0.114,0.078]	[-0.089,0.089]	[-0.11,0.078]
Treatment dummy	0.204	0.333	0.203
	[0.038,0.365]	[0.155,0.48]	[0.037,0.362]
Log child's cognitive skill (t)	0.130	0.067	0.133
	[0.016,0.246]	[-0.044,0.18]	[0.017,0.25]
Log child's socio-emotional skill (t)	-0.032	0.021	-0.033
	[-0.133,0.087]	[-0.083,0.145]	[-0.131,0.088]
Log mother's cognitive skill (t)	0.754	0.367	0.755
	[0.582,0.939]	[0.162,0.498]	[0.583,0.943]
Log mother's socio-emotional skill (t)	0.071	0.024	0.071
	[-0.008,0.139]	[-0.06,0.108]	[-0.008,0.139]
Log number of children (t)	-0.129	-0.134	-0.128
	[-0.18,-0.077]	[-0.186,-0.072]	[-0.18,-0.078]
Toy price	-0.095	-0.020	-0.094
	[-0.168,-0.027]	[-0.085,0.037]	[-0.163,-0.026]
Food price	0.096	0.042	0.097
	[0.006,0.178]	[-0.026,0.121]	[0.006,0.178]
Maternal childhood exposure to conflict	-0.011	-0.096	
	[-0.08,0.063]	[-0.139,-0.032]	
Rank test[(a)]	0.011		
Cragg-Donald test	0.019		
Test of joint significance - p-values:			
Toy price, food price, conflict	0.028	0.008	
Toy price, food price, conflict, treatment	0.001	0.001	
Toy price, food price	0.013	0.455	0.010
Toy price, food price, treatment	0.001	0.000	0.000
Conflict, treatment	0.047	0.000	

Note: Dependent variables are the log of material and time investments. Maternal childhood exposure to conflict is defined as the number of conflicts against civil population/1000 when the mother was a child. t refers to baseline/pre-treatment measurement. [(a)] This is a test of the null hypothesis that the smallest eigenvalue of the 2×2 matrix $\beta'\beta$ is zero, where β is the 3×2 matrix of coefficients on food price, toy price and conflict in the material and time investment equations. We present this alongside the Cragg-Donald test because in this context it is not clear which is the more powerful. 95% Confidence intervals in square brackets. These as well as the p-values for the rank tests and all other tests are based on 1,000 bootstrap replications of the entire estimation process taking into of clustering at the municipality level.

and particularly so for the former; however mother's socio-emotional skills only affect material investments significantly and there the impact is very small. The number of other children in the household at baseline reduce significantly both time and material investments: an extra child reduces both investments by about 13%, which is consistent with a quantity/quality tradeoff among children. Moreover, the results are in line with a model where parents choose investments taking into account complementarity with child cognitive skills.

Importance of instruments. The next set of variables reported in Table 5 explain investments and act as excluded instruments when we estimate the production functions and treat investments as endogenous. These are the prices of toys and food and the level of conflict in the municipality when the mother was a child. The rank test we implement has a p-value of 0.011.[30] As an alternative, we also present the Cragg-Donald form of this test, which has a p-value of 0.019. These establish the strength of the instruments, allowing for the fact that there are two endogenous variables.

The prices we consider are expected to affect material investment through the household budget constraint. Indeed, the p-value of a test of joint significance of these instruments is 0.013 in the material investment equation. As we would expect, toy price has a negative and significant impact on material investments. Food price, on the other hand, has a significant positive effect on material investment, implying that play materials and food are substitutes.

Understanding the determinants of quality time is harder, as argued above. Prices have no explanatory power: their joint p-value in the quality time investment equation is 0.46. In results not reported here, we also found that wages did not predict time investments either, supporting our argument that the opportunity cost of quality time is not the time spent in the labor market and that spending quality time with children

[30] See Robin and Smith (2000) and Blundell et al. (1998).

does not necessarily require monetary resources, but rather perhaps knowledge of child development and a certain willingness to engage with the child. Maternal childhood exposure to conflict, on the other hand, has a strong and significant negative impact on quality time activities (p-value 0), although it has no impact on material investments. The latter result reinforces the idea that this variable is not merely picking up some omitted background factor.

Jointly, these instruments are strong enough to allow us to control for the endogeneity of investments in the production function. In some specifications, we also exclude treatment status from the production functions, thus implicitly using the treatment as an instrument. In this case, the instruments only become stronger.

Finally, in the last column of Table 5 we present a specification for material investment that excludes the conflict variable; the coefficients are almost identical to those in column (1), and the prices are jointly even more significant. We use this investment equation to estimate a model where only material investments enter the production function for cognitive skills.

6.2 The production function for cognitive skills

In Table 6, we report estimates of a Cobb-Douglas production function for cognitive skills. The first column presents results where investments are assumed to be conditionally exogenous; in the remaining columns investments are taken as endogenous.

The production function demonstrates a high level of persistence for cognition; however, socio-emotional skills do not affect cognition at this early age, and this result remains unchanged whether we treat investments as endogenous or not. Mothers' cognition and socio-emotional skills have a strong positive effect on cognitive development when we use OLS, but in the remaining columns, when we use IV for the investments, these effects disappear, implying that mothers' skills operate through the initial conditions of the child and through the investment decisions only. Finally, the

Table 6: Estimates of the production function for cognitive skills

	(1)	(2)	(3)	(4)	(5)
	OLS	IV			
Instruments:		*Prices, conflict*	*Price, conflict, treatment*	*Prices*	*Prices and treatment*
TFP	-0.016	-0.02	0.008	-0.006	-0.009
	[-0.094,0.053]	[-0.111,0.085]	[-0.088,0.104]	[-0.092,0.072]	[-0.083,0.057]
TFP * Treatment	0.084	0.076		-0.022	
	[-0.025,0.192]	[-0.12,0.391]		[-0.176,0.161]	
Log child's cognitive skill (t)	0.67	0.663	0.634	0.632	0.631
	[0.589,0.77]	[0.544,0.797]	[0.521,0.748]	[0.527,0.747]	[0.533,0.739]
Log child's socio-emotional skill (t)	-0.004	0.004	0.019	0.014	0.02
	[-0.091,0.087]	[-0.09,0.142]	[-0.095,0.142]	[-0.088,0.129]	[-0.081,0.129]
Log mother's cognitive skill (t)	0.217	0.018	-0.179	-0.099	-0.094
	[0.089,0.35]	[-0.435,0.477]	[-0.548,0.198]	[-0.503,0.291]	[-0.45,0.21]
Log mother's socio-emotional skill (t)	0.106	0.089	0.064	0.067	0.074
	[0.031,0.173]	[-0.031,0.162]	[-0.034,0.149]	[-0.021,0.157]	[-0.025,0.154]
Log of number of children (t)	0.041	0.075	0.084	0.09	0.086
	[-0.01,0.092]	[-0.068,0.15]	[0.011,0.163]	[0.004,0.176]	[0.023,0.163]
Log of material investment (t+1)	0.08	0.463	0.799	0.531	0.516
	[0.016,0.157]	[0.033,1.182]	[0.197,1.4]	[0.038,1.015]	[0.195,0.949]
Log of time investment (t+1)	0.04	-0.131	-0.323		
	[-0.051,0.129]	[-1.178,0.3]	[-1.022,0.227]		
Gap in output between treated and control:					
(a) Measured in the data			0.115		
(b) Predicted by the model estimates	0.114	0.127	0.055	0.086	0.105

Note: Dependent variable is the log cognition of the child at follow up. t and $t+1$ denote that a variable was measured at baseline and endline, respectively. 95% Confidence intervals in brackets based on 1,000 bootstrap replications of the entire estimation process taking into account clustering at the municipality level.

number of other children in the family improve child cognitive development. This is particularly interesting because the investment equations show that additional children reduce both material and quality time investments. However the presence of other (for the most part older) children has a direct impact on child development, with an extra child improving cognition by 4%-9% depending on the specification.

We now turn to the estimates of the direct treatment effect and the coefficients on parental investments, as measured by our 'material investment' and our 'quality

time investment' factors. When we treat investments as exogenous and use OLS, we find that the impact of material investments on cognitive development is significantly different from zero, but that of time is insignificant. The direct treatment effect is large but very imprecisely estimated, to the extent that it is insignificant. The estimate of the impact of material investments increases dramatically when we treat investment as endogenous. Going from column 1 to column 2, this coefficient increases from 0.08 to 0.46; and although it is now estimated less precisely it remains significant at the 5% level. The change in the size of the coefficient points to parents compensating for negative shocks affecting their children.

The coefficient on time investment, instead, stays insignificant and its point estimate actually turns negative. The estimate of the direct effect of the intervention stays insignificant, although, as we discussed above it is strongly significant in the investment equations. For this reason, in column 3, we force the direct treatment effect to be zero. Despite the now unequivocal strength of the instruments (since we are also using the treatment allocation as an exclusion restriction in column 3), the results do not change: material investments enter with a large and significant coefficient (0.56) and time investments remain completely insignificant.

Using material investments alone Given the measurement issues for quality time, which is based on self-reports and discussed further below,[31] and in light of the results presented thus far, in the next two columns we exclude time investments from the production function and we switch to using *only* prices as instruments so the relevant first stage investment equation is in the third column of Table 5. In one specification we include the treatment dummy reflecting the direct effect of the intervention (column 4) and in the other we exclude it (column 5). The results are essentially the same, though with improved precision, and present a clear message: material investments

[31] See also section 4.2.

have a strong and positive impact, while the direct effect of treatment is small and completely insignificant.

Interpreting the impact of the intervention Based on these estimates, we now consider how the intervention affects outcomes through the lens of the production function. The possible channels include changes in the production function - a direct effect - and changes in parental investment inputs. Changes in the production function could happen for a number of reasons. First, the weekly session of the home visitor with the child, as well as any materials left behind, can be thought of as a new input; second the intervention could lead to a better use of measured inputs by parents or equivalently an improvement in the unmeasured quality of these inputs. These are possible channels through which the intervention could affect outcomes over and above inducing more investments through its emphasis on parenting and the direct involvement of the mother in the home visit.

From the coefficients reported in column 1 of Table 6, which treats investment as exogenous, together with the increase in investment documented in Table 5, we see that increased investment accounts for about 25% of the impact of the intervention. The point estimate of the coefficient on treatment, instead, indicates that the intervention directly increased cognition considerably, by about 8.4%, although it is not precisely estimated. Therefore, according to this specification, parental investments play some role in mediating the intervention but there is a large direct effect, at least in terms of the point estimate.

A key problem with these results, though, is that they assume investments are invariant to unobserved shocks to child development that occur between baseline and the end of the intervention 18 months later. The IV results presented in columns 2 to 5 of Table 6 address this issue and present a different story: the increase in material investments now explains a large fraction of the observed impact. Thus, a

good description of cognitive development among children aged 3-4 is that it is driven by initial cognition (measured at aged 1-2), stimulation provided by older siblings, and material investments provided by parents, and the main channel through which the intervention affects cognitive development is by shifting these investments. The model in column 5 with material investments alone achieves a good fit of the data and captures over 91% of the overall impact of the intervention. Finally, we find no evidence that any other coefficient of the production function changed. In Table E.2 we allow all coefficients to vary by treatment status. The overall p-value is 0.32, while for the non-intercept coefficients it is 0.29.

These results highlight the importance of material investments but leave no role for quality time investments. This is perhaps surprising because one would expect quality time to be an important input as well. However, this result could at least partly reflect the measurement issue we discussed earlier (see Section 4.2) and that needs to be addressed by future research into data collection: material investments are measured through interviewer observations, while time investments are self-reported and hence carry the risk to be over-reported, particularly in the intervention communities where the importance of quality time has been continuously emphasized throughout the duration of the intervention. Material investments provided by the parents require real effort through making or buying toys, and they are likely to correlate strongly with actual activities carried out with the child. For example, providing a toy, whether it is home made or bought, will likely include spending time with the child in this activity. The actual observation of play materials may therefore better reveal the developmental activities carried out by parents. On the other hand the materials left behind by the intervention and captured by the intervention/treatment dummy may not reflect actual engagement by the parent. This again emphasizes that the program works to the extent that it shifts actual parental behavior.

6.3 The production function for socio-emotional skills

In Table 7, we present estimates of several specifications of the production function for socio-emotional skills. As with cognitive skills, we observe that the accumulation process of socio-emotional skills exhibits substantial amount of persistence (regardless of the specification considered). The point estimates, however, are lower than in the case of cognition: about 0.50, compared to about 0.70 for cognitive skills. The intervention also has no direct effect in any specification. The lagged value of cognitive development is marginally significant in the various specifications, although the size of the coefficient is not very large.

Table 7: Estimates of the production function for socio-emotional skills

	(1)	(2)	(3)	(4)	(5)
	OLS	IV			
Instruments:		*Prices, conflict*	*Prices, conflict*	*Prices, conflict, treatment*	*Prices, conflict, treatment*
TFP	-0.008	-0.007	0.000	-0.026	-0.022
	[-0.08,0.063]	[-0.09,0.065]	[-0.083,0.066]	[-0.095,0.058]	[-0.092,0.051]
TFP * Treatment	-0.012	-0.083	-0.102		
	[-0.124,0.093]	[-0.318,0.158]	[-0.308,0.123]		
Log child's cognitive skill (t)	0.105	0.086	0.087	0.104	0.098
	[0.018,0.192]	[-0.018,0.211]	[-0.014,0.188]	[0.005,0.22]	[0.004,0.191]
Log child's socio-emotional skill (t)	0.51	0.494	0.5	0.512	0.513
	[0.403,0.659]	[0.38,0.669]	[0.388,0.662]	[0.388,0.665]	[0.404,0.656]
Log mother's cognitive skill (t)	-0.073	-0.091	-0.143	-0.035	-0.083
	[-0.217,0.049]	[-0.438,0.305]	[-0.361,0.085]	[-0.293,0.329]	[-0.23,0.075]
Log mother's socio-emotional skill (t)	0.036	0.041	0.06	0.046	0.044
	[-0.058,0.119]	[-0.055,0.144]	[-0.054,0.135]	[-0.047,0.134]	[-0.041,0.126]
Log of number of children (t)	0.098	0.128	0.131	0.103	0.101
	[0.047,0.153]	[0.015,0.235]	[0.03,0.223]	[0.026,0.17]	[0.036,0.164]
Log of material investment (t+1)	0.142	-0.079		-0.117	
	[0.06,0.256]	[-0.645,0.424]		[-0.698,0.341]	
Log of time investment (t+1)	0.119	0.558	0.542	0.436	0.324
	[-0.006,0.213]	[-0.175,1.285]	[-0.054,1.15]	[-0.096,1.022]	[0.023,0.687]
Gap in output between treated and control:					
(a) Measured in the data			0.087		
(b) Predicted by the model estimates	0.057	0.087	0.078	0.121	0.108

Note: Dependent variable is the log socio-emotional skills of the child at follow up. t and $t+1$ denote that a variable was measured at baseline and endline, respectively. 95% Confidence intervals in brackets based on 1,000 bootstrap replications of the entire estimation process - clustering at the municipality level.

Parental background variables, such as mothers' cognitive and socio-emotional skills, do not seem to matter for the development of children's socio-emotional skills. Instead, the number of siblings has a positive and significant impact on these skills, consistent with what we found for cognitive development. This result is robust across specifications.

As with cognitive skills, we experiment with a number of specifications. In none of the specifications does the direct treatment effect play any role: the coefficient is always negative and very imprecise. When we use OLS we find that material investments matter, and time investments enter with a marginally significant coefficient. When we turn to IV (columns 2 to 5), none of the investment coefficients remain significant, with the exception of the results in column 5. There, we exclude material investments, which always enter with a negative coefficient when instrumented, and the treatment dummy, which never plays a role. In this case the results suggest that our measure of quality time can explain the improvement in socio-emotional skills.

Based on the OLS and using a specification that excludes the direct impact of the intervention (which is negative and insignificant) all the reported coefficients in column 1 remain unchanged. That specification explains 79% of the impact (it predicts a 0.069 log point improvement). The alternative specification in column 5 over predicts the impact. Whether we take the OLS results or the IV results presented in column 5, their implications are clear and similar: the intervention acted by improving parental engagement with the children and had not direct impact on their development.

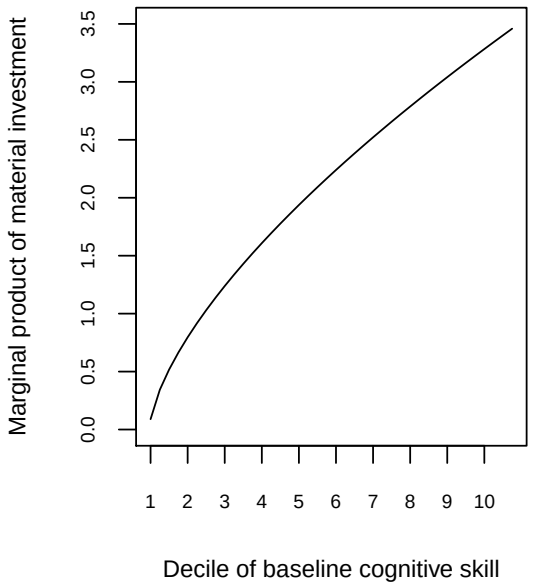
6.4 Complementarities between inputs of the production function

The Cobb-Douglas specification implied by the data means that the inputs are complementary (Cunha et al., 2006; Heckman and Mosso, 2014). The complementarity is imposed by the functional form. However, in earlier versions of the paper, when we estimated the substitution elasticity, we always found it very close to 1. Given the

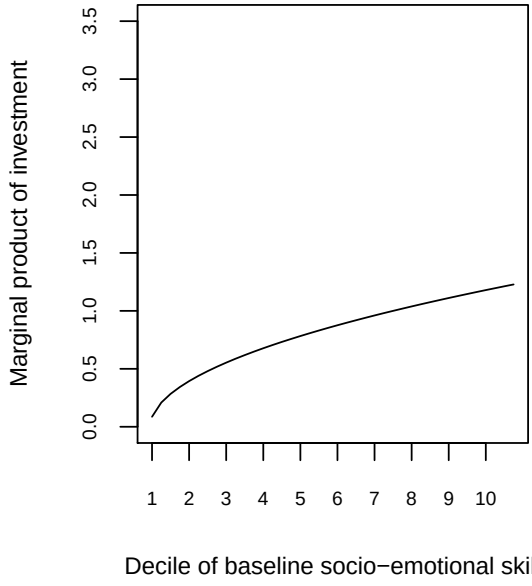
metric we use for the latent factors, the return to investment is higher for children with better initial conditions. In Figure 2(a) we show the effect of one standard deviation increase in material investment on cognitive skills and in Figure 2(b) the effect of one standard deviation increase in both material and time investments on socio-emotional skills. The y-axes are in standard deviation units of the outcome.

Figure 2: Complementarity between investments and baseline skills

(a) Production function for cognitive skills



(b) Production function for socio-emotional skills



Notes: Figure 2.a (2.b) is based on the estimates of the production function for cognitive skills (socio-emotional skills) reported in Column 5 (1) of Tables 6 and 7. The figures above are constructed by evaluating the increase in cognitive (socio-emotional skills) in standard deviation units resulting from an increase in one standard deviation of investments at different deciles of $\theta^C_{i,t}$ for (a) and $\theta^S_{i,t}$ for (b) and holding all remaining inputs of the production function at their mean values across the sample.

The complementarity of investments with initial conditions of the child may appear contradictory to the set of studies indicating that early interventions benefit low-achieving children the most (Bitler et al., 2014; Elango et al., 2016). However, one needs to allow for the differences in the populations concerned. Our intervention targets the 20% poorest children in Colombia. While these children do not live in extreme poverty, they may still be poorer and of lower ability at baseline than disadvantaged

children targeted by programs such as Head Start in the US. Our results imply that, in this subset of the population, those with a better start benefited more. However, one can imagine that with a population that extends more broadly in the socio-economic distribution, diminishing returns could set in unless perhaps we design an intervention better attuned to higher ability children.

6.5 Implications for longer term outcomes

The results have implications for what to expect in the longer term. Under the assumption that the patterns of self-productivity and complementarities documented here remain the same at least in the medium-term, the fact that the return to investments are complementary to the prior level of achievement implies that, if parents keep investing at the higher levels induced by the intervention, the skills of the treated children should continue to improve, subject of course to the impact of investments at later ages. And since ability and investments are complementary, future investments would further increase the skills of the intervention group. However, the estimates also reveal fadeout between the baseline and the follow up: the coefficient on past cognitive (socio-emotional) skills is about 0.7 (0.5) in the production function for cognitive (socio-emotional) skill. If such mean reversion continues beyond the ages that we consider and if parents revert to the level of activity in the control group we can expect the impact of the intervention to become much smaller in the long-run. In fact, as shown in Andrew et al. (2018), parents in our intervention reduced their level of engagement to that of the control group when interviewed again two years after the end of the intervention. And consistent with the results here the impact of the intervention also faded.

Partial fadeout was also observed following the Jamaica intervention (Walker et al., 2005), although perhaps because the original effect was as large as 80% of a standard deviation, half the original impact remained. This underscores two key lessons. First,

we should not underestimate the challenges involved in scaling up successful small scale efficacy trials, such as the Jamaica study, and in achieving comparable initial impacts. Second, we need to better understand what motivates parents to continue investing in children in the longer run and whether sustained intervention is needed to preserve and reinforce initial gains in such environments.

7 Conclusion

Children from poor backgrounds accumulate developmental deficits from a very early age. Causes include not only the risky environments in which they live but also the lack of stimulation, which prevents the brain from developing its full potential. Such adverse early experiences are at the heart of the intergenerational transmission of poverty.

In this paper, we present results from an early childhood intervention carried out in Colombia that promoted suitable parenting and stimulation to children between one and three and a half years old. The intervention involved weekly home visits delivered by local women who had no prior specific knowledge of child development, but were trained as part of the intervention to deliver a structured stimulation curriculum that progressed in difficulty. The evaluation by randomized controlled trial showed improvements in a number of developmental dimensions, including cognition, language and socio-emotional development, though impacts on the latter are smaller.

We use data from the experiment to estimate a model of parental investments in children and production functions for children's cognition and socio-emotional skills. The aim is to improve our understanding of the development of child skills from a very early age and to provide an interpretation of how the intervention affected child development. The model estimates trace some of the origins of social inequalities to the beginning of life: children with higher initial skills obtain more investments from their parents, and, given their skills, mothers with higher levels of cognition invest more in

their children.

The estimates of the production functions also provide evidence of several important features of skill development among children below the age of 4. First, we find strong evidence of self-productivity of skills. That is, the current stock of cognitive (socio-emotional) skills strongly affects the development of future cognitive (socio-emotional) skills, but also imply mean reversion. Second, we find evidence of cross-productivity: the current stock of cognitive skills fosters the development of future socio-emotional skills (although the effect is small), but the reverse does not seem to be the case. This result contrasts with that reported by Cunha, Heckman, and Schennach (2010), who find socio-emotional skills to be important for the accumulation of future cognitive skills though at a different age. While the presence of siblings in the household reduces parental investments, it improves the cognitive and socio-emotional development of our subject child (who is the youngest child in the family in most cases), most likely through interactions and imitation, over and beyond the effect of parental investments.

Most importantly for the question addressed in this paper, as well as more broadly, our results show that investments help develop both cognitive and socio-emotional skills. The program increased investments substantially and our key conclusion is that it is this increase that led to the estimated impacts of the intervention on children's skills. Specifically, the intervention increased the cognitive development of the children by 0.115 log points and socio-emotional development by 0.087 log points. Our best estimates of the production functions imply that the increase in parental investments induced by the program account for just over 91% of the intervention impact on cognition and at least 66% of its impact on socio-emotional skills. There was no direct impact of the intervention. It is thus the parenting component of the intervention – where the home visitor directly involves the mother in developing the stimulation activities – that underlies its success.

Our study answers some important questions but raises many more, calling for

further experimentation and analysis. We need to understand how to better target and treat the most disadvantaged of society. Moreover, the analysis raises the question of how sustainable the effects of the intervention are and how salient improvements at this age are for longer term outcomes. This requires longer-term follow ups of the children participating in the intervention and calls for further research with systematic measurements and interventions at various stages of life. Finally, an important lesson from our results is that involving parents in interventions is key to promoting child development in the short-term. Going forward however, it is crucial that we better understand how to incentivize and ensure continued parental investments after the intervention has ended.

References

Agostinelli, F. and Wiswall, M. (2016). Identification of dynamic latent factor models: The implications of re-normalization in a model of child development. NBER Working Paper No. 22441.

Almlund, M., Duckworth, A., Heckman, J. J., and Kautz, T. (2011). Personality, psychology and economics. In E. Hanushek, S. Machin, and L. Woessmann, eds., *Handbook of the Economics of Education*. Elsevier Science.

Almond, D. and Currie, J. (2011). Human capital development before age five. In O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics, Volume 4b*. Elsevier Science, Amsterdam.

Anderson, T. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 5, 111–150. University of California Press, Berkeley.

Andrew, A., Attanasio, O., Fitzsimons, E., Grantham-McGregor, S., Meghir, C., and Rubio-Codina, M. (2018). Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in colombia. *PLoS Med* 15(4): e1002556. Https://doi. org/10.1371/journal.pmed.1002556.

Attanasio, O., Fernandez, C., Fitzsimons, E., Grantham-McGregor, S., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastucture of a Conditional Cash Transfer programme to deliver a scalable integrated early child development programme in Colombia: A cluster randomised controlled trial. *British Medical Journal* 349:g5785.

Attanasio, O., Fitzsimons, E., Gomez, A., Gutiez, M.-I., Meghir, C., and Mesnard, A. (2010). Children's schooling and work in the presence of a conditional cash transfer program in rural colombia. *Economic Development and Cultural Change* 181–210.

Attanasio, O., Meghir, C., and Nix, E. (2017). Investments in children and the development of cognition and health in india. NBER Working Paper No. 21740.

Attanasio, O., Meghir, C., and Santiago, A. (2012). Education choices in Mexico: Using a structural model and a randomized experiment to evaluate Progresa. *Review of Economic Studies* 79(1):37–66.

Bates, J., Freeland, C., and Lounsbury, M. (1979). Measurement of infant difficultness. *Child Development* 50:794–803.

Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development*. Harcourt Assessment, San Antonio, TX, 3rd edition.

Betancourt, T. S. (2015). Intergenerational effects of war. *JAMA Psychiatry* 72(3):199–200.

Bitler, M., Hoynes, H., and Domina, T. (2014). Experimental evidence on distributional impacts of head start. NBER Working Paper 20434.

Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., McCoy, D. C., Fink, G., Shawar, Y. R., Shiffman, P. J., Devercelli, A. E., Wodon, Q. T., Vargas-Bar., and Grantham-McGregor, S. (2016). Early childhood development coming of age: science through the life course. *Lancet. 6736* 1–14.

Blundell, R., Duncan, A., and Meghir, C. (1998). Estimating labor supply responses using tax reforms. *Econometrica* 66(4):827–761.

Caldwell, B. and Bradley, R. (2001). *HOME Inventory and Administration Manual*. University of Arkansas for Medical Sciences, Little Rock, AK, third edition edition.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science* 343:1478 – 1485. DOI: 10.1126/science.1248429.

Campbell, F. A. and Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development* 65:684 – 698. Doi:10.2307/1131410 Medline.

Carneiro, P., Hansen, K., and Heckman, J. (2003). Estimating distribution of counterfactuals with an application to the returns to schooling and measurement of the effect of uncertainty on schooling choice. *International Economic Review* 44(2):361–422.

Cattell, R. (1966). The scree test for the number of factors. *Multivariate Behavioral Research* 1:629–637.

Creech, S. K. and Misca, G. (2017). Parenting with ptsd: A review of research on the influence of ptsd on parent-child functioning in military and veteran families. *Frontiers in Psychology* 8. Doi:10.3389/fpsyg.2017.01101.

Cunha, F. and Heckman, J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43(4):738–782.

Cunha, F., Heckman, J., Lochner, L., and Masterov, D. (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek, S. Machin, and L. Woessmann, eds., *Handbook of the Economics of Education, Volume 1*. Elsevier Science, Amsterdam.

Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the technology of cognitive and non-cognitive skill formation. *Econometrica* 78(3):883–931.

DelBoca, D., Flinn, C., and Wiswall, M. (2014). Household choices and child development. *Review of Economic Studies* 81:137–185.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review* 102(4):1241–78.

Dunn, L., Padilla, E., Lugo, D., and Dunn, L. (1986). *Manual del Examinador para el Test de Vocabulario en Imenes Peabody (Peabody Picture Vocabulary Test), Adaptaci Hispanoamericana (Hispanic-American Adaption)*. Pearson Assessments, Minneapolis, MS.

Elango, S., García, J., Heckman, J., and Hojman, A. (2016). Early childhood education. volume Economics of Means-Tested Transfer Programs in the United States, volume 2. University of Chicago Press, Chicago, IL.

Fernald, L., Kariger, P., Engle, P., and Raikes, A. (2009). *Examining child development in low-income countries: A toolkit for the assessment of children in the first five years of life*. The World Bank, Washington DC.

Frongillo, E., Sywulka, S., and Kariger, P. (2003). UNICEF psychosocial care indicators project. Final report to UNICEF. Mimeo, Cornell University.

Gelber, A. and Isen, A. (2010). Children's schooling and parents' behavior: Evidence from the head start impact study. *Journal of Public Economics* 101:25 ? 38.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeerch, C., Walker, S., Chang, S., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* 344:998–1001.

Gorsuch, R. (1983). *Factor Analysis*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Gorsuch, R. (2003). Factor analysis. In *Handbook of psychology: Research methods in psychology*, volume 2, 143–164. John Wiely & Sons, Inc., Hoboken, NJ.

Grantham-McGregor, S., Cheung, Y., Cueto, S., Glewwe, P., Richter, L., and Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *The Lancet* 369(9555):60–70.

Grantham-McGregor, S., Powell, C., Walker, S., and Himes, J. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study. *Lancet* 338(758):1–5.

Hamadani, J., Tofail, F., Hilaly, A., Huda, S., Engel, P., and Grantham-McGregor, S. (2010). Use of Family Care Indicators and their relationship with child development in bangladesh. *Journal of Health Population and Nutrition* 28(1):23–33.

Heckman, J., Moon, S., Pinto, R., Savelyev, P., and Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Journal of Quantitative Economics* 1:1–46.

Heckman, J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Reviews of Economics* 6:689–733.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review* 103:2052–2086.

Helmers, C. and Patnam, M. (2011). The formation and evolution of childhood skill acquisition: Evidence from India. *Journal of Development Economics* 95(2):252–266.

Horn, J. (1965). A rationale and a test for the number of factors in factor analysis. *Psychometrika* 30:179–185.

Hu, Y. and Schennach, S. (2008). Instrumental variable treatment of nonclassical measurement error models. *Econometrica* .

Jackson-Maldonado, D. (2011). Spanish cdi-iii development and preliminary results. Presentation at the International Congress for the Study of Child Language, Montreal.

Jackson-Maldonado, D. and Conboy, B. (2011). Preliminary norms and validity for the Spanish CDI-III. American Speech-Language-Hearing Association (ASHA) Annual Conference. San Diego, CA.

Jackson-Maldonado, D., Marchman, V., and Fernald, L. (2012). Short form versions of the Spanish MacArthur-Bates Communicative Development Inventories. In *Applied Psycholinguistics*. Cambridge University Press, Cambridge.

Jackson-Maldonado, D., Thal, D., Marchman, V., Newton, T., Fenson, L., and Conboy, B. (2003). *MacArthur Inventarios del Desarrollo de Habilidades Comunicativas: User's guide and technical manual*. Brookes, Baltimore, MD.

Kaiser, H. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20:141–151.

Knudsen, E. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience* 16(1):1412–1425.

Knudsen, E., Heckman, J., Cameron, J., and Shonkoff, J. (2006). Economic, neurobiological and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Science* 103(27):10155–10162.

Lancet (2016). Early child development in developing countries 2016 .

Lu, C., Black, M. M., and Richter, L. M. (2016). Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *Lancet Global Health 4* e916?e922.

Nielsen, E. R. (2015). Achievement gap estimates and deviations from cardinal com- parability. Finance and Economics Discussion Series 2015-040. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.040.

Putnam, S. P., Gartstein, M., and Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior questionnaire. *Infant Behavior and Development* 29(3):386–401.

Putnam, S. P., Jacobs, J., Garstein, M., and Rothbart, M. (2010). Development and assessment of short and very short forms of the Early Childhood Behavior Questionnaire. International Conference on Infant Studies. Baltimore, MD.

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

Raven, J. (1981). Manual for Raven's progressive matrices and vocabulary scales. In *Research Supplement No.1: The 1979 British Standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, Together With Comparative Data From Earlier Studies in the UK, US, Canada, Germany and Ireland*. Harcourt Assessment, San Antonio, TX.

Robin, J. and Smith, R. (2000). Tests of rank. *Econometric Theory* 16:151–175.

Romano, J. P. and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica* 73:1237 – 1282.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica* 39(3):577–591.

Rubio-Codina, M., Attanasio, O., Meghir, C., Varela, N., and Grantham-McGregor, S. (2015). The socioeconomic gradient of child development: Cross-sectional evidence from children 6-42 months in bogota. *Journal of Human Resources* 50(2):464–483.

Schennach, S. (2004). Estimation of nonlinear models with measurement error. *Econometrica* 72:33–75.

Schennach, S. (2007). Instrumental variable estimation of nonlinear errors-in-variables models. *Econometrica* 75:201–239.

Todd, P. and Wolpin, K. (2006). Using experimental data to validate a dynamic behavioral model of child schooling and fertility: Assessing the impact of a school subsidy program in Mexico. *American Economic Review* 96(5):1384–1417.

Velicer, W. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika* 41:32l–327.

Walker, S., Chang, S., C.Powell, E.Simonoff, and Grantham-McGregor, S. (2006). Effects of psychosocial stimulation and dietary supplementation in early childhood on psychosocial functioning in late adolescence: follow-up of randomized controlled trial. *British Medical Journal* 333(7566):472.

Walker, S. P., Chang, S. M., Powell, C. A., and Grantham-McGregor, S. M. (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *The Lancet* 366:1804 – 1807.

Walker, S. P., Wachs, T. D., Grantham-McGregor, S. M., Black, M., Nelson, C., and Huffman, S. (2011). Inequality in early childhood: risk and protective factors for early child development. *Lancet* 378(1325-38).