# Estimating the production function for human capital: Results from a randomized control trial in Colombia

O. Attanasio (UCL/IFS), S. Cattan (IFS), E. Fitzsimons (IFS),
C. Meghir (Yale/IFS), and M. Rubio-Codina (IFS)

Trinity College Dublin - January 21 2014

# Using a conditional cash transfer programme to scale up an integrated early child development intervention in Colombia: a cluster randomised controlled trial

O. Attanasio (UCL- IFS), C. Fernandez (Mathematica), E. Fitzsimons (IFS),
S. Grantham-McGregor (UCL), C. Meghir (Yale-IFS), and M. Rubio-Codina (IFS)

EDePo @ IFS
Evaluation of Development Policies at the Institute for Fiscal Studies

# Outline

# Introduction

- Substantial body of research on the development of human capital
  - Almond and Currie (2011)
  - Cunha et al. (2006), Cunha and Heckman (2008), Cunha et al. (2010)
  - The Lancet Series (2007, 2010)
  - Developing countries: Attanasio et al. (2013), Grantham-McGregor et al. (2012), Helmers and Patnam (2011)...

# Introduction

- Substantial body of research on the development of human capital
  - Almond and Currie (2011)
  - Cunha et al. (2006), Cunha and Heckman (2008), Cunha et al. (2010)
  - The Lancet Series (2007, 2010)
  - Developing countries: Attanasio et al. (2013), Grantham-McGregor et al. (2012), Helmers and Patnam (2011)...
- Early human capital plays an important role for adult outcomes

# Introduction

- Substantial body of research on the development of human capital
  - Almond and Currie (2011)
  - Cunha et al. (2006), Cunha and Heckman (2008), Cunha et al. (2010)
  - The Lancet Series (2007, 2010)
  - Developing countries: Attanasio et al. (2013), Grantham-McGregor et al. (2012), Helmers and Patnam (2011)...
- Early human capital plays an important role for adult outcomes
- Reducing gaps in early skills can help reduce social and economic inequality

# Introduction

- Substantial body of research on the development of human capital
  - Almond and Currie (2011)
  - Cunha et al. (2006), Cunha and Heckman (2008), Cunha et al. (2010)
  - The Lancet Series (2007, 2010)
  - Developing countries: Attanasio et al. (2013), Grantham-McGregor et al. (2012), Helmers and Patnam (2011)...
- Early human capital plays an important role for adult outcomes
- Reducing gaps in early skills can help reduce social and economic inequality

# Introduction

- This evidence raises some important questions:
  - ⇒ How does human capital develop?
  - ⇒ What role, if any, can policy play to remedy early deficiencies among children?
  - ⇒ What kind policies are effective at scale?

# Introduction

- This evidence raises some important questions:
  - $\Rightarrow$ How does human capital develop?
  - $\Rightarrow$ What role, if any, can policy play to remedy early deficiencies among children?
  - $\Rightarrow$ What kind policies are effective at scale?

- Human capital formation is a complex process
  - Human capital is multi-dimensional (cognitive, non-cognitive, health...)
  - Skill formation is a dynamic process
  - Dimensions of human capital interact both within and across periods
  - Both genes and the environment are important inputs

# The importance of the early years for policy

- The early years is a particularly salient period for policy
  - Human capital is malleable (and vulnerable)
  - Dynamic complementarities ("skills beget skills")

- Well-designed and well-targeted interventions in the early years can partially compensate for exposure to adverse environments

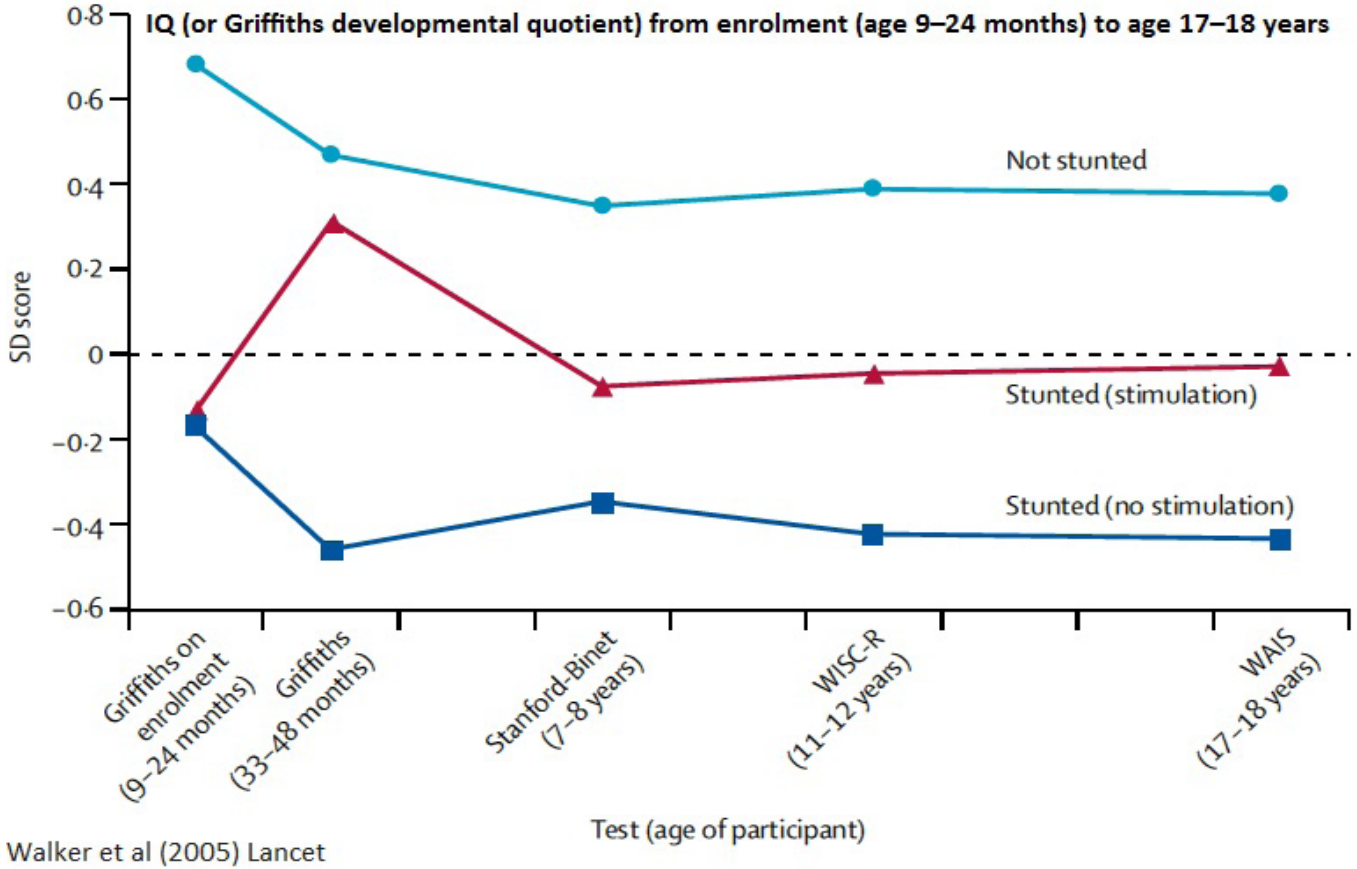
- Prominent studies have demonstrated strong results sustained in the long-run
  - Perry School experiment (Anderson (2008); Heckman et al. (2010, 2011, 2013))
  - Abecedarian Project (Mass and Barnett, 2002)
  - Jamaica Study (Walker et al. (1990, 2011); Gertler et al. (2013))

# The Jamaica study

- The Jamaica experiment included three treatments and a control group
  - The treatments were:
  - Infant Stimulation
  - Nutrition (calories)
  - Both
- The stimulation followed a structured curriculum, that we will discuss later
- It was delivered by professional health assistants
- It targeted children from 9-24 months and the intervention lasted 2 years

# The Jamaica study

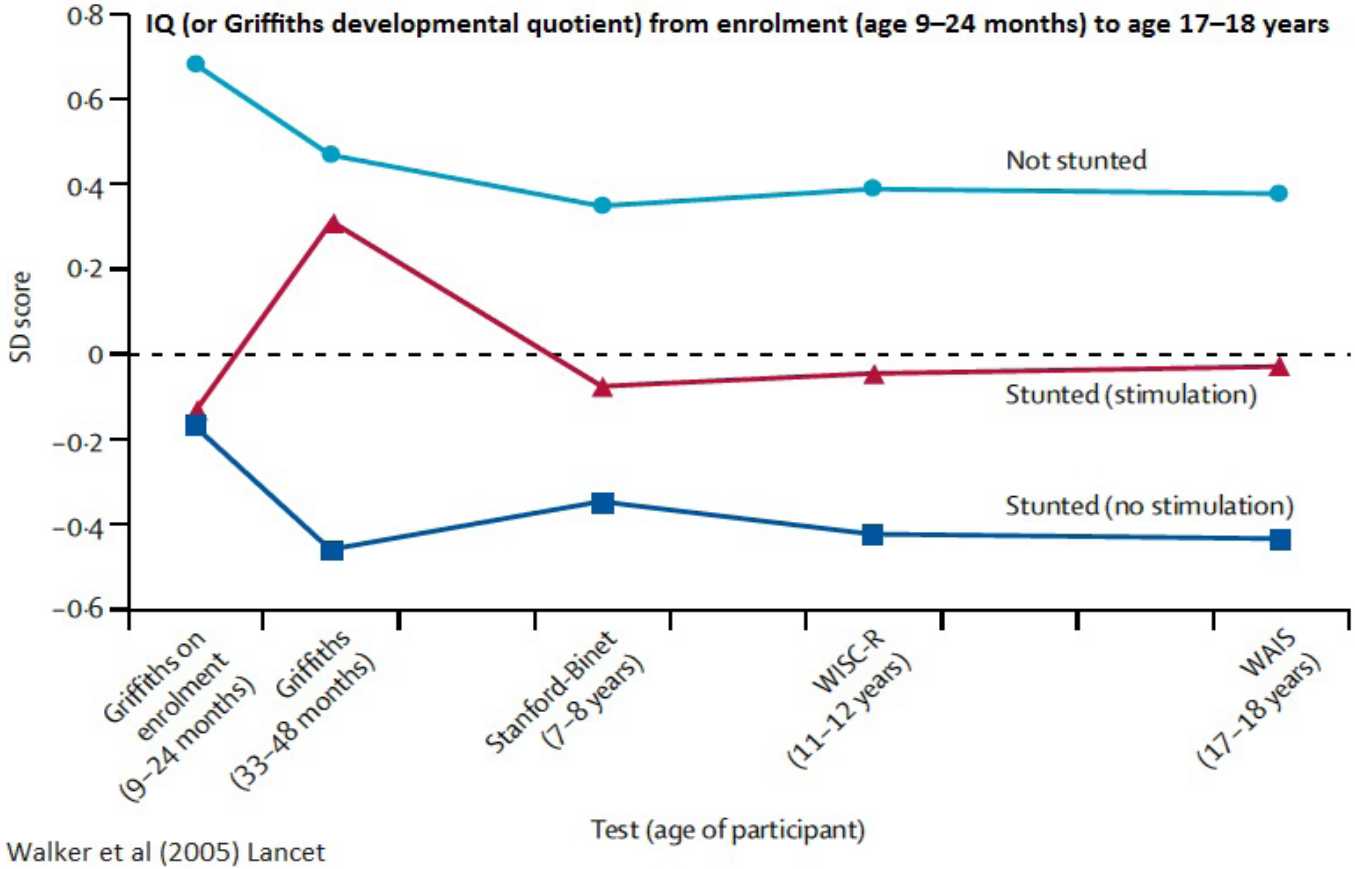
- Grantham-McGregor and colleagues have demonstrated using the Jamaica experiment that cognition effects are sustainable




Walker et al (2005) Lancet

# The Jamaica study

- Grantham-McGregor and colleagues have demonstrated using the Jamaica experiment that cognition effects are sustainable




Walker et al (2005) Lancet

- Recently Gertler, Heckman, McGregor et al. (2012) have shown that the effects are as important in labor market outcomes.

# Outline

# The intervention in Colombia

- In this context, we designed and implemented an early childhood intervention in Colombia
  - The basic structure was guided by the Jamaica experiment by Sally Grantham-McGregor et al. 1991 - Lancet (SGM)
  - However there are two important new elements:
- Intervention: the emphasis on designing the program using local resources in a scalable fashion
- Research Design: collect detailed household data to allow modeling the behavioral impact of the intervention to identify mechanisms

# The Intervention

- Rather than using professional health workers, we select local women to implement the intervention.

# The Intervention

- Rather than using professional health workers, we select local women to implement the intervention.
- We target our intervention to the beneficiaries of Familias en Accion - a CCT program.
  - The target population belong to the lowest economic group in terms of poverty as classifies by the SISBEN system

# The Intervention

- Rather than using professional health workers, we select local women to implement the intervention.
- We target our intervention to the beneficiaries of Familias en Accion - a CCT program.
  - The target population belong to the lowest economic group in terms of poverty as classifies by the SISBEN system
- This group is represented by elected women - Madres Lideres (MLs)
  - The MLs are better educated, more pro-active but still they are part of the community they are intended to serve.

# The Intervention

- Rather than using professional health workers, we select local women to implement the intervention.
- We target our intervention to the beneficiaries of Familias en Accion - a CCT program.
  - The target population belong to the lowest economic group in terms of poverty as classifies by the SISBEN system
- This group is represented by elected women - Madres Lideres (MLs)
  - The MLs are better educated, more pro-active but still they are part of the community they are intended to serve.
- This is the key element for the scalability of the program.

# Scalability

- Using local representatives has a number of advantages:
  - The intervention costs are low
  - The local women may become agents of change within their communities
  - The communities may take ownership of the intervention thus making it sustainable.

# The Intervention design

- We adapted the Jamaica curriculum to the Colombian context.
  - We trained 6 professionals, each was assigned to 8 villages.
  - Our professionals (supervisors) trained 3/4 'madre lideres' in each village.
  - The MLs were trained for three weeks.

# The Intervention design

- MLs were hired on a part time basis by us.

# The Intervention design

- MLs were hired on a part time basis by us.
- A scaled up intervention could do better and would have to have a regular update to the training

# The Intervention design

- MLs were hired on a part time basis by us.
- A scaled up intervention could do better and would have to have a regular update to the training
- After training, the supervisors kept going to the villages on a regular basis:
  - monitoring the implementation, giving feedback and counseling

# The Intervention design

- MLs were hired on a part time basis by us.
- A scaled up intervention could do better and would have to have a regular update to the training
- After training, the supervisors kept going to the villages on a regular basis:
  - monitoring the implementation, giving feedback and counseling
- The monitors/ supervisors were in constantly in touch with the MLs sent them motivational messages and short information.

# The Intervention design

- Each ML visited 5-6 children and their mothers and distributed the micronutrients.
  - weekly visits of one hour each.
- The intervention lasted for 18 months.
  - Two years would probably be better but we had inadequate funds
- The intervention is cheap:
  - US$500 per year per child.
  - 50% of cost is monitoring and supervision.
  - At scale it can be reduced to US$300 US.

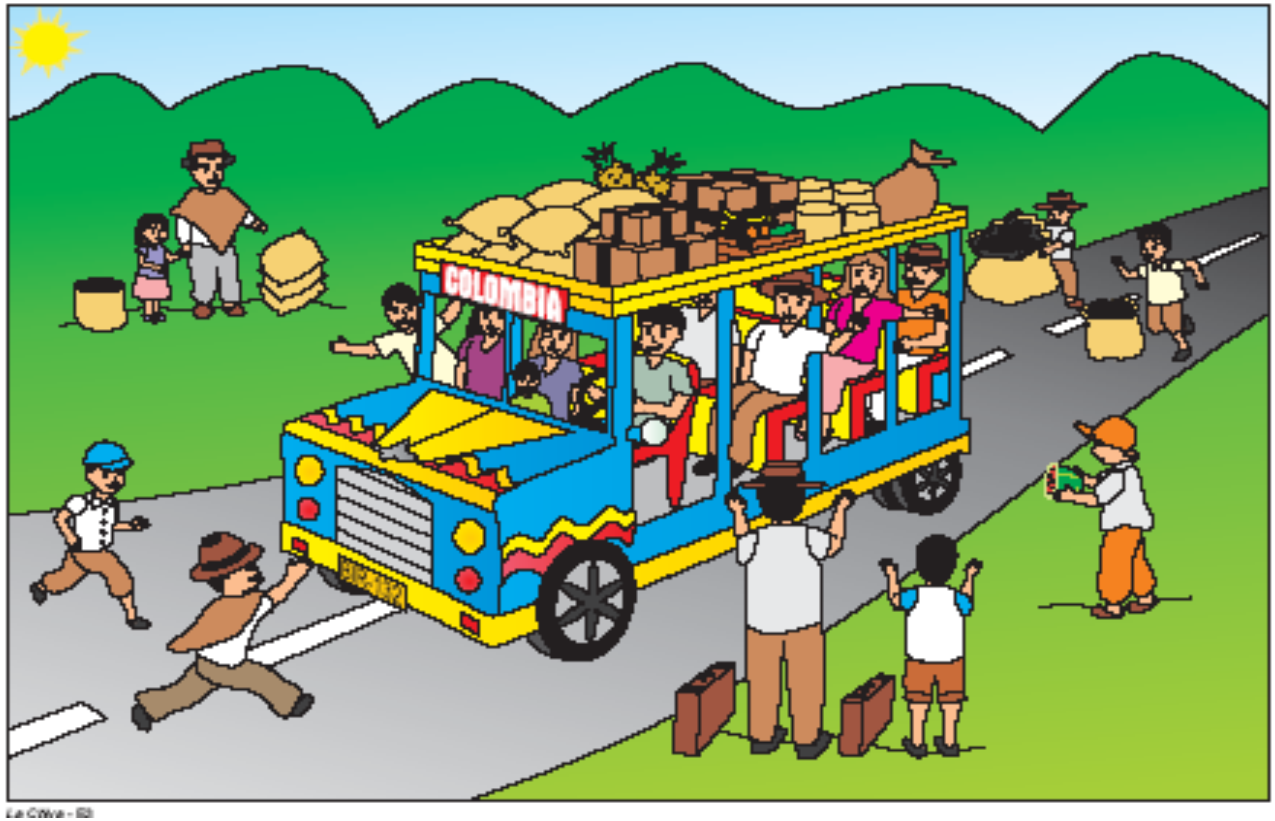
# The curriculum

- Promote child-development in an integrated manner:
  - motor, language, cognitive, socio-emotional
- Encourage mothers to teach her children based on events surrounding daily routine activities
- Involve other children or members of the family where possible  this could generate important spillovers.

# The curriculum

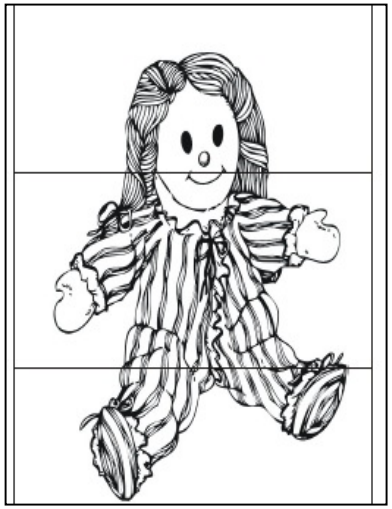
- Picture Books
- Pictures to stimulate conversation
- Puzzles
- Cubes/Blocks and patterns
- Toys from recycled material
- Language games and songs.

# The curriculum

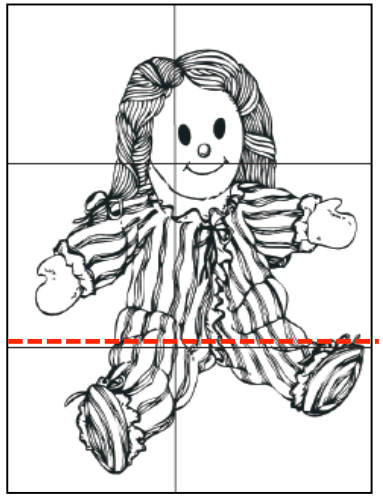


# The curriculum

Rompecabezas Pallaso
(21 meses en adelante)



Rompecabezas Muñeca
3 piezas (31 meses +)
6 piezas (41 meses +)



# The curriculum

# The curriculum

# Summary of Research Questions

- At some level we understand that well designed ECD interventions can produce spectacular results

# Summary of Research Questions

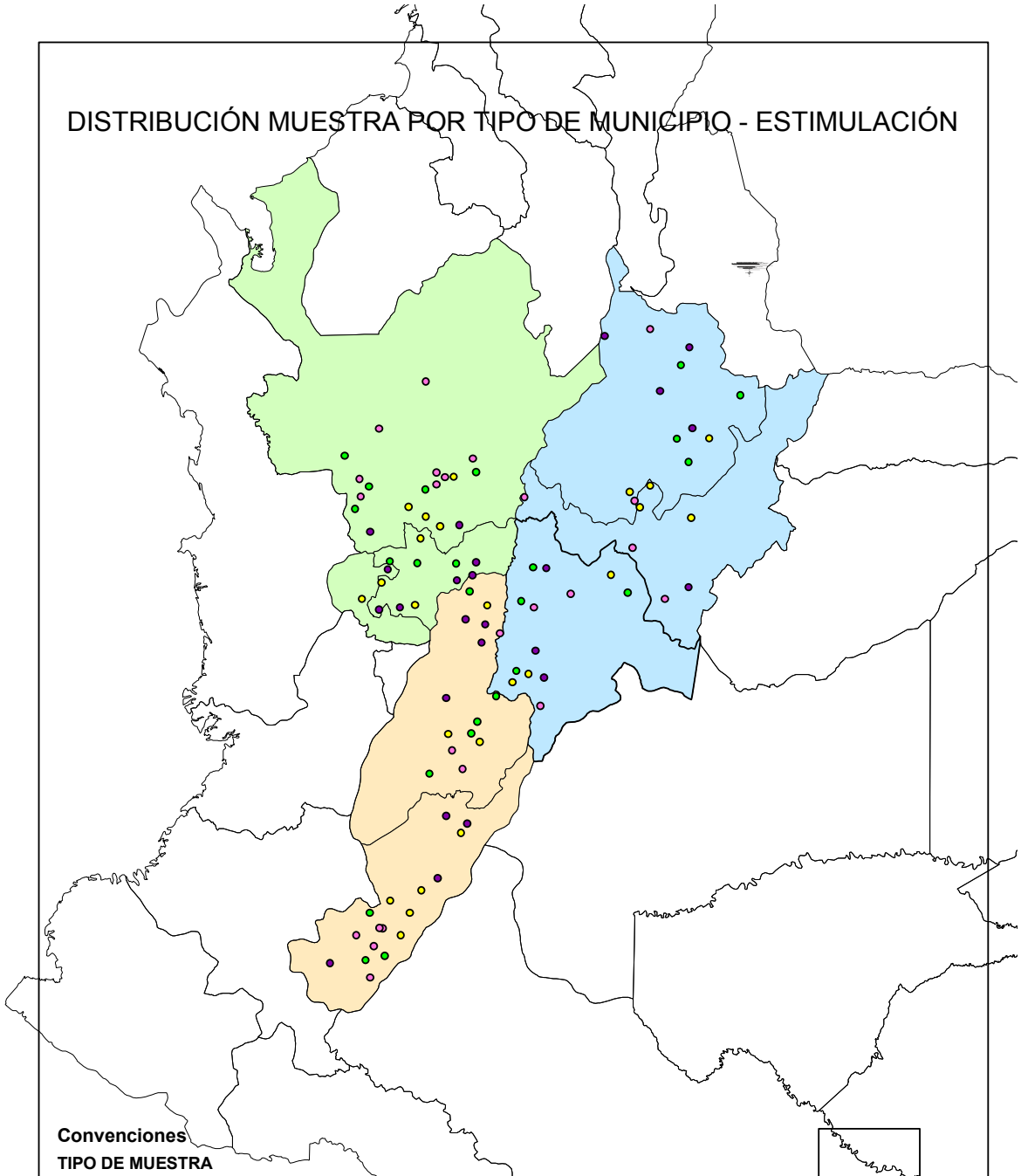
- At some level we understand that well designed ECD interventions can produce spectacular results
- Here we pose to new research questions:
  - Can we make it work by drawing on local resources?
  - Why do these interventions work? How do households change their behavior? What is the HC production function and how does it change?
  - How do the effects vary by economic environment, gender etc.?

# Outline

# Research design

- It targeted 1,429 children aged 12-24 months at baseline in 96 semi-urban towns
- Children were randomized into 4 groups (at the village-level)
  1. Stimulation
  2. Supplementation (micronutrients)
  3. Stimulation + Supplementation
  4. Control

# The Random Assignment



# Evaluation

- Choosing the children/families:
- In both treatment and control we drew randomly (3+2) 5 MLs
- The families with children in the 1-2 year age group became our subject families (in both treatment and control)
- If the ML refused to participate we still kept the families so there is no selection bias between treatment and control. We just replaced the ML and kept the same families

# Measurement

- Rich data collected on human capital and investments
  - Baseline: 12 - 24 months old
  - Follow-up I: 30 - 42 months old
  - Follow-up II: 54 - 66 months old (to be collected this Fall)

# Measurement tools

- Many measurement on children development
  1. Motor and Cognitive Development: Bailey Test
  2. Socio-emotional Development: Bates Temperament
  3. Language Development: MacArthur-Bates
  4. Height, weight, haemoglobin and Morbidity
  5. Food Intakes (target child and ¡6 children in household)
  6. Child care arrangements and Time Use (target child and ¡6 children in household)

# Measurement tools

- Mothers and families
  - general household survey
  - Education , labour supply and time use
  - Reproductive history, Health conditions, Depression.
  - Health Condition
  - Height, weight and haemoglobin
  - Aversion to Inequality and to Risk
  - Depression (CESD)
  - Knowledge on Parenting
  - Parenting Practices and the Home Environment

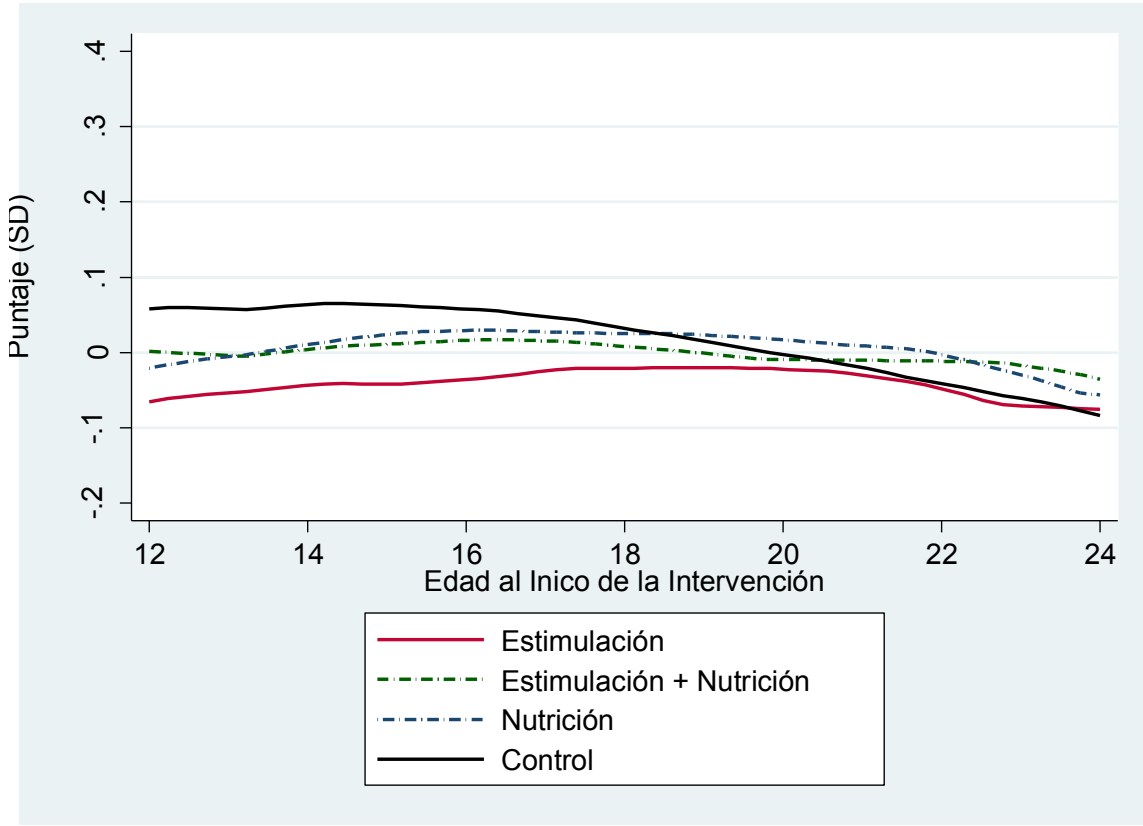
# Measurement tools

- HomeVisitor questionnaire
  - Education, labour supply and time use
  - Health Condition
  - Aversion to Inequality and to Risk
  - Knowledge on Parenting and Children

# Measurement tools

- HomeVisitor questionnaire
  - Education, labour supply and time use
  - Health Condition
  - Aversion to Inequality and to Risk
  - Knowledge on Parenting and Children
- Process data
- Focus groups

# Experimental balance



# Baseline Results

## Mother's Health

| Mother's variables | | Home Stimulation | Nutrition + Home Stimulation | Nutrition | Control |
|---|---|---|---|---|---|
| Anaemia | Mean | 23.42% | 19.71% | 22.29% | 20.12% |
| | Std. Error | 2.32% | 2.16% | 2.29% | 2.23% |
| BMI | Mean | 25.29 | 24.99 | 24.59 | 25.16 |
| | Std. Error | 0.27 | 0.26 | 0.24 | 0.26 |
| CESD 10 | Mean | 8.41 | 8.91 | 9.57 | 9.36 |
| | Std. Error | 0.30 | 0.29 | 0.30 | 0.31 |
| Depressed | Mean | 38.44% | 42.35% | 44.88% | 46.44% |
| | Std. Error | 2.67% | 2.68% | 2.73% | 2.78% |

# Baseline Results

## Child Health

| Z-scores | | Control | Home Stimulation | Nutrition + Home Stimulation | Nutrition |
|---|---|---|---|---|---|
| Weight for age | Mean | -0.26 | -0.16 | -0.18 | -0.17 |
| | Std. Error | 0.05 | 0.06 | 0.05 | 0.05 |
| Lenght (height) for age | Mean | -0.83 | -0.73 | -0.74 | -0.58* |
| | Std. Error | 0.06 | 0.06 | 0.06 | 0.06 |
| BMI for age | Mean | 0.35 | 0.39 | 0.37 | 0.25 |
| | Std. Error | 0.05 | 0.06 | 0.05 | 0.05 |
| weight/lenght for age | Mean | 0.20 | 0.25 | 0.23 | 0.15 |
| | Std. Error | 0.05 | 0.06 | 0.05 | 0.05 |

- There are clear nutritional deficiencies
- Substantial stunting relative to international standard
- Height deficiency, but BMI above international standard

# Baseline Results

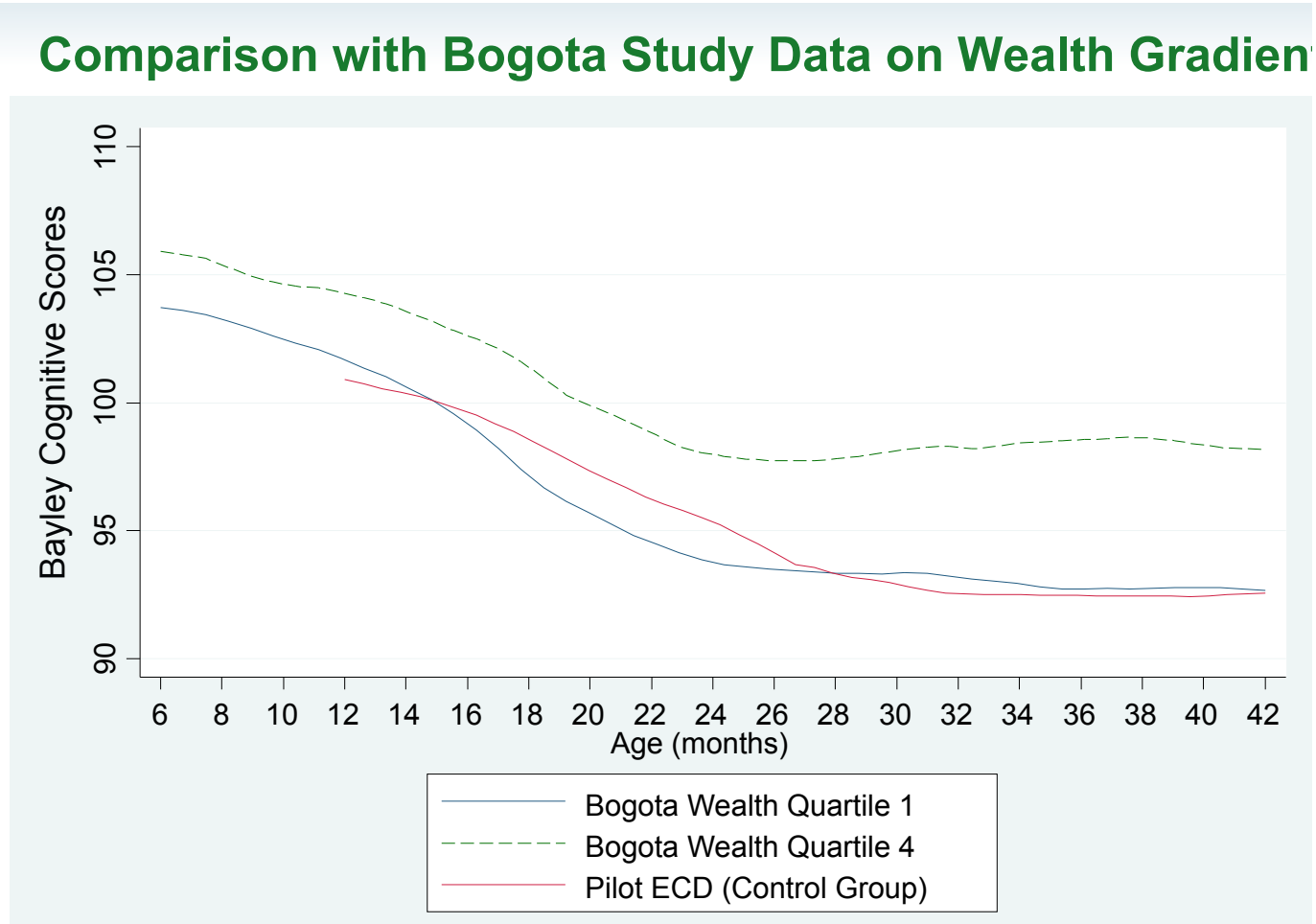
## Child's Health

| Variable | | Control | Home Stimulation | Nutrition + Home Stimulation | Nutrition |
|---|---|---|---|---|---|
| Diarrhea (last 15 days) | Mean | 37.01% | 34.47% | 33.33% | 37.93% |
| | Std. Error | 2.64% | 2.54% | 2.51% | 2.60% |
| Stunting | Mean | 12.24% | 11.11% | 10.73% | 8.33% |
| | Std. Error | 1.79% | 1.68% | 1.65% | 1.48% |
| Wasting | Mean | 0.60% | 2.84%* | 1.41% | 1.15% |
| | Std. Error | 0.42% | 0.89% | 0.63% | 0.57% |
| Undernourishing | Mean | 2.69% | 4.56% | 3.95% | 1.44% |
| | Std. Error | 0.88% | 1.11% | 1.04% | 0.64% |
| Anaemia | Mean | 41.49% | 41.31% | 38.98% | 41.67% |
| | Std. Error | 2.70% | 2.63% | 2.60% | 2.65% |

# Baseline Results

| | Congnitive | Language | Motor |
|---|---|---|---|
| | b/se | b/se | b/se |
| Gender (==1 if male) | -0.872 | -4.429*** | -1.057 |
| | (0.58) | (0.7) | (0.76) |
| Age | -0.698*** | -0.346** | 0.444*** |
| | (0.1) | (0.11) | (0.11) |
| Kid is anaemic | -1.395* | -0.945 | -0.522 |
| | (0.57) | (0.65) | (0.83) |
| Kid is wasted | -1.707 | -1.68 | -3.189 |
| | (2.94) | (2.72) | (3.33) |
| Kid is stunted | -2.945** | -3.007** | -2.914* |
| | (0.94) | (1.08) | (1.35) |
| kid is undernourished | -2.952 | -0.999 | -3.539 |
| | (2.76) | (2.48) | (2.97) |
| Years of education Mother | 0.325*** | 0.430*** | 0.228* |
| | (0.09) | (0.09) | (0.1) |
| Factor: Utilities | 0.572 | 0.208 | -0.028 |
| | (0.31) | (0.41) | (0.41) |
| Factor: Possesions | 0.468* | 0.427 | 0.664** |
| | (0.18) | (0.25) | (0.23) |
| Number of children under 6 years | -1.022* | -2.507*** | -1.246** |
| | (0.44) | (0.52) | (0.43) |
| Weight at birth | 0 | 0 | 0 |
| | (0) | (0) | (0) |
| Was he breastfed? | 4.187* | 0.953 | 3.392 |
| | (2.1) | (1.78) | (2.58) |
| constant | 101.981*** | 99.237*** | 83.833*** |
| | (3.78) | (4.63) | (5.33) |

# Wealth Gap - Age and Cognition




Wealth Gap.pdf

# Attrition

- Sample Loss between household survey and Bayley test Baseline: 9 children (0.62%).
- Attrition between survey rounds (18 months): Household Survey: 3.52%.

# Spatial Dependence and Precision

- The design consists overall of 24 communities in each branch and about 15 children per community
- It was quite hard to predict spatial correlation in advance given the kind of outcomes we were considering
- It turns out that the spatial correlation once we condition on baseline characteristics is down to about 0.04 or less (depending on the outcome).
- So this implies an effective sample size of about 220 per variant (880 overall)
- This implies that our study has much larger effective sample size than the Jamaica study (for example) where the total sample size was 129 (32 observations per variant)

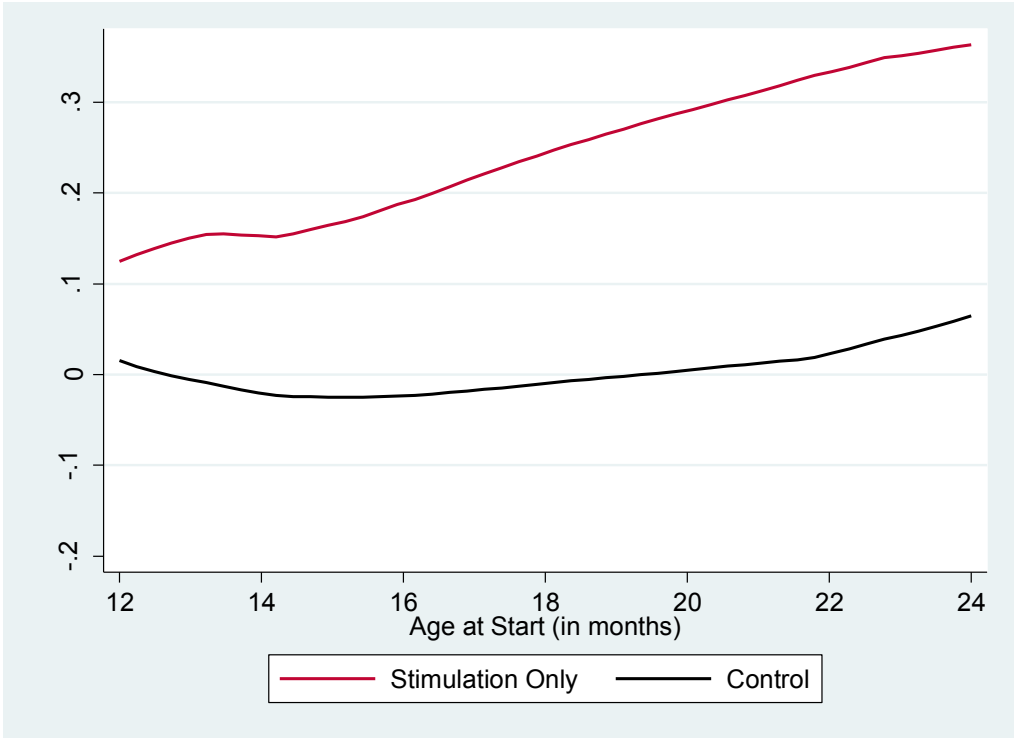
# Outline

# Some impacts

## Effects on Cognition (Bayley) by Age at Intervention Start

| | All | 12-18 mths | 18-24 mths |
|---|---|---|---|
| Stim (in SD) | 0.251** | 0.197+ | 0.357** |
| | (0.073) | (0.111) | (0.092) |

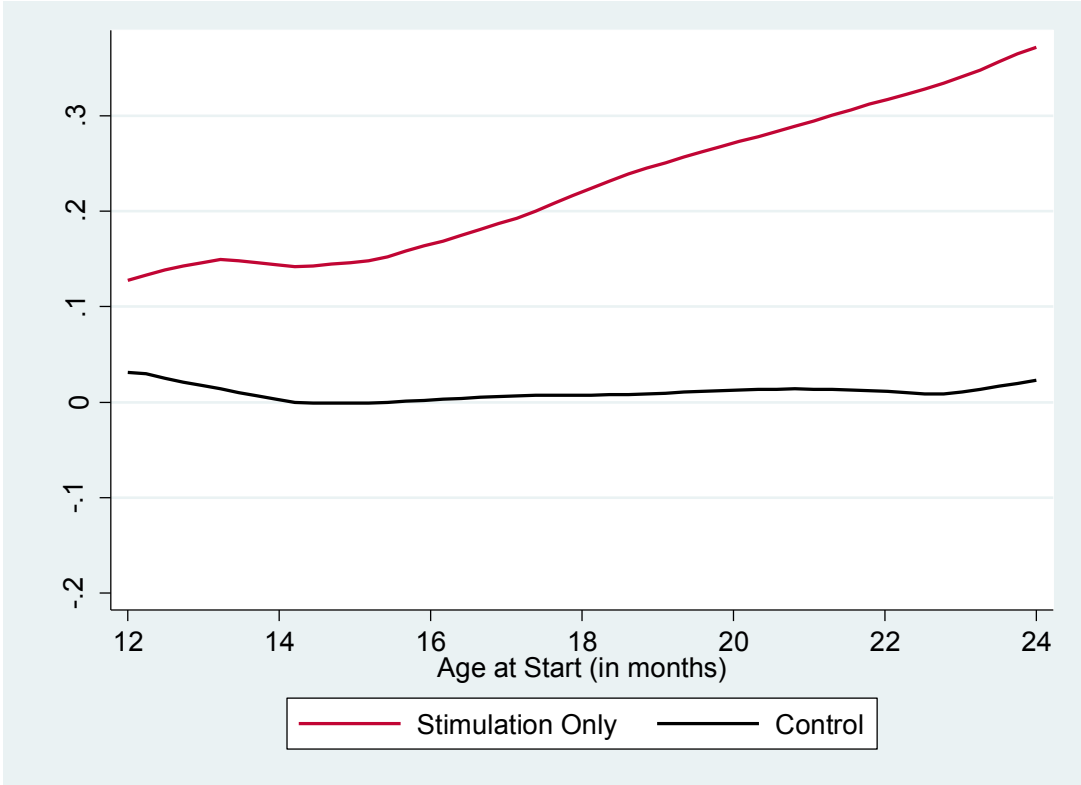
n = 1267; +significant at 10%, *significant at 5%, **significant at 1%



# Effects on Receptive Language (Bayley) by Age at Start

| | All | 12-18 mths | 18-24 mths |
|---|---|---|---|
| Stim (in SD) | 0.188** | 0.174 | 0.258* |
| | (0.080) | (0.113) | (0.120) |

n =1267; +significant at 10%, *significant at 5%, **significant at 1%



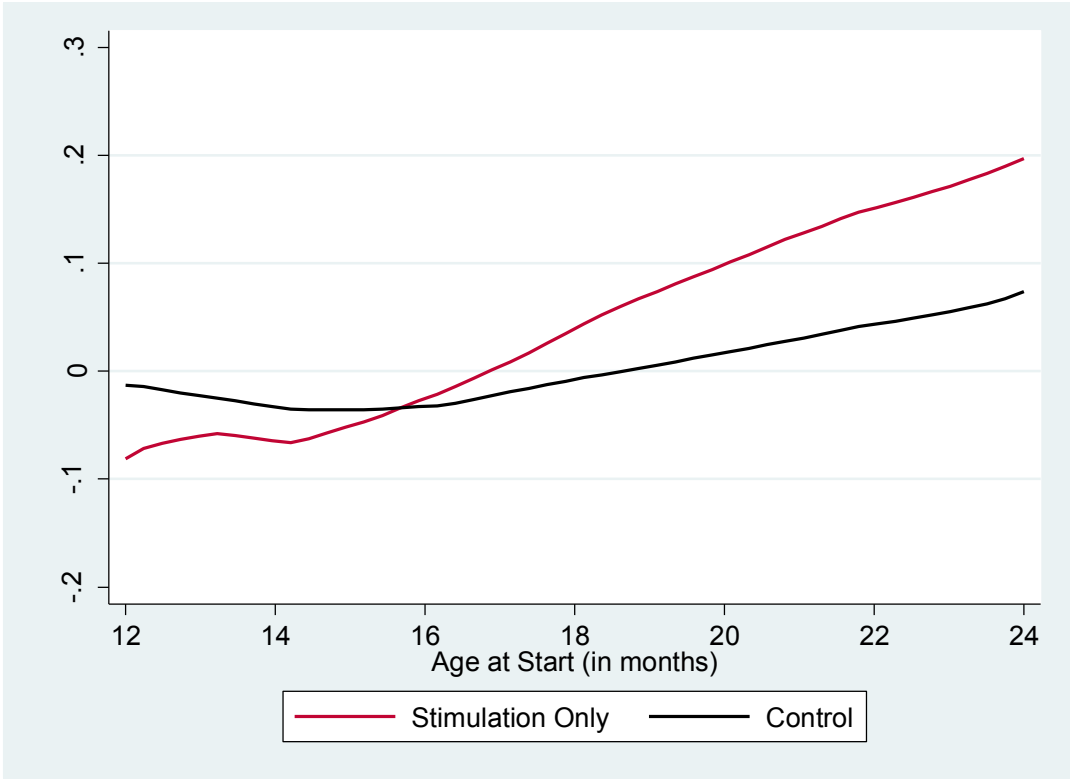

Stim - Lang Rec.pdf

# Some impacts

## Effects on Expressive Language (Bayley) by Age at Start

| | All | 12-18 mths | 18-24 mths |
|---|---|---|---|
| Stim (in SD) | 0.059 | 0.023 | 0.191 |
| | (0.073) | (0.122) | (0.159) |

n =1267; +significant at 10%, *significant at 5%, **significant at 1%



# Summary of evaluation findings

Effects of the treatments on cognitive and non-cognitive measures (in standard deviations)

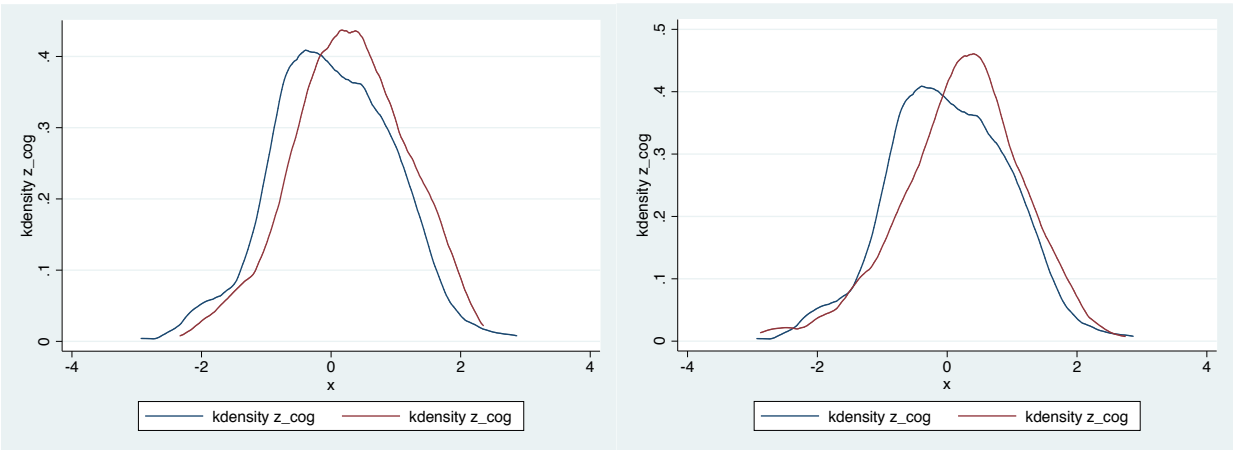
| Outcome: | Cognition | Receptive Language | Expressive Language | Number of Words Can Say | Child is difficult |
|---|---|---|---|---|---|
| *Scale:* | *(Bayley)* | *(Bayley)* | *(Bayley)* | *(MacArthur)* | *(Bates)* |
| Stim | 0.251** | 0.188** | 0.059 | 0.147$^{+}$ | -0.127$^{+}$ |
| | (0.073) | (0.080) | (0.073) | (0.080) | (0.067) |
| Stim + Supp | 0.206** | 0.162* | 0.079 | 0.171* | -0.037 |
| | (0.071) | (0.073) | (0.080) | (0.085) | (0.059) |
| Supp | 0.047 | 0.039 | 0.084 | 0.130$^{+}$ | -0.014 |
| | (0.059) | (0.084) | (0.087) | (0.076) | (0.062) |
| N | 1,267 | 1,267 | 1,267 | 1,325 | 1,325 |

Note: $^{+}$significant at 10%, *significant at 5%, **significant at 1%

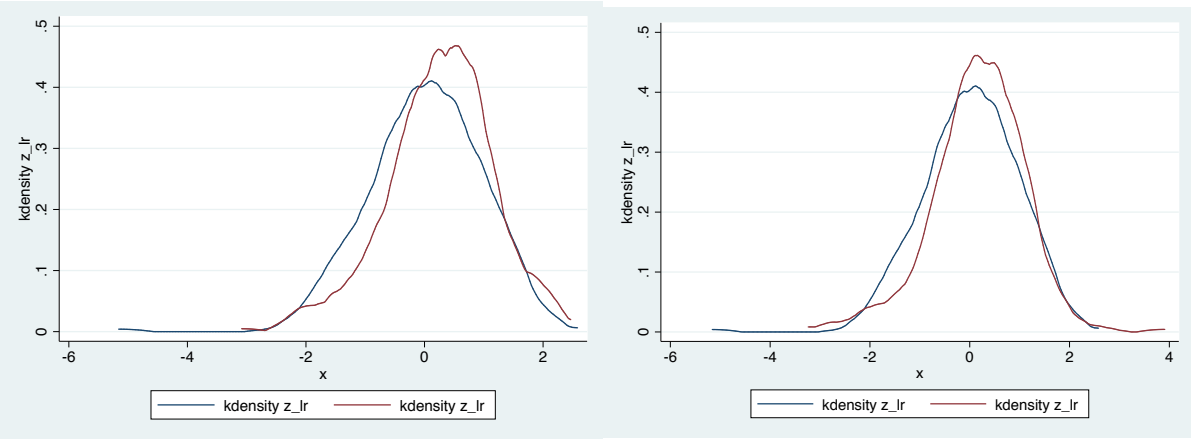
$\Rightarrow$ Home visits had large impacts on child's cognitive and language development

Even larger impacts for older children (.35 for cognition, .27 for language)

# Impacts along the distribution - cognition



# Impacts along the distribution - Receptive Language




recept-Stim-Stimnutr.pdf

# Some suggestive evidence that the intervention changed parental behavior

| | Varieties of Play Materials | Number of Play Activities | Reading/ Looking Picture bks | Telling Stories | Naming/ Counting |
|---|---|---|---|---|---|
| Stim | 0.556** | 0.564** | 0.202** | 0.098* | 0.093** |
| | (0.128) | (0.152) | (0.039) | (0.041) | (0.032) |
| Stim + Supp | 0.452** | 0.731** | 0.169** | 0.079$^{+}$ | 0.139** |
| | (0.137) | (0.153) | (0.038) | (0.041) | (0.034) |
| Supp | 0.213 | 0.217 | -0.015 | 0.058 | 0.005 |
| | (0.167) | (0.153) | (0.038) | (0.036) | (0.039) |
| *Mean outcome* | *3.715* | *0.153* | *0.314* | *0.262* | *0.553* |

# Outline

# The production function of human capital

- These results raise the question of how the intervention produced such effects
- To answer this question, we need a framework to understand the determinants of child development
- In this paper, we:
  - Estimate a production function for child development
  - Use this framework to understand how the intervention worked
- The exercise is similar in spirit to Heckman, Pinto and Savelyev (2013).

# The production function of human capital

- These results raise the question of how the intervention produced such effects
- To answer this question, we need a framework to understand the determinants of child development
- In this paper, we:
  - Estimate a production function for child development
  - Use this framework to understand how the intervention worked
- The exercise is similar in spirit to Heckman, Pinto and Savelyev (2013).
- There are many empirical challenges to estimating production functions
- We use the approach of Cunha, Heckman and Schennach (2010) and extend it to study the impact of the intervention

# A general model of child development

- Each child starts with a particular endowment of skills $\theta_1 = (\theta_{c,1}, \theta_{n,1})$ at the baseline age.
  - These initial conditions can be influenced by family environment and genetics
  - We focus on cognitive ($k = c$) and non-cognitive skills ($k = n$)

- There are several periods in childhood $t = 1, \ldots, T$

- We describe the formation of human capital with a CES production function

$$\theta_{k,t+1} = A[\gamma_{k,1}\theta_{c,t}^{\rho_k} + \gamma_{k,2}\theta_{n,t}^{\rho_k} + \gamma_{k,3}P_c^{\rho_k} + \gamma_{k,4}P_n^{\rho_k} + \gamma_{k,5}I_{t+1}^{\rho_k}]^{\frac{1}{\rho_k}} e^{\eta_{k,t}}$$

  - Current stock in each dimension of human capital $(\theta_{c,t}, \theta_{n,t})$
  - Mother's cognitive and non-cognitive skills $P = (P_c, P_n)$
  - Investments $(I_t)$
  - Unobserved shocks $(\eta_t)$

- The role of these factors may vary with age and stages of development
  - For now, we only focus on one stage (1 to 3.5 years old)

# Incorporating the role of the intervention in the framework

- We consider the impact of any stimulation ($s = 1$) vs. no stimulation ($s = 0$)
- The intervention could have affected the formation of skills:
  - By shifting the distribution of investments
  - By shifting the productivity of the inputs
- The production function we will estimate is:

$$\theta_{k,t+1} = A_s[\gamma_{k,1}\theta_{c,t}^{\rho_k} + \gamma_{k,2}\theta_{n,t}^{\rho_k} + \gamma_{k,3}P_c^{\rho_k} + \gamma_{k,4}P_n^{\rho_k} + \gamma_{k,5}I_{t+1}^{\rho_k}]^{\frac{1}{\rho_k}} e^{\eta_{k,t}}$$

where the joint distribution of factors $(\theta_{t+1}, \theta_t, I_{t+1}, P)$ can differ between treated and controls

# Outline

# Table of Contents

# Measuring skills and investments

- At baseline and follow-up, we collected measurements of skills and investments
- These measurements come in many forms
  - Psychological instruments (which usually include many items per (sub)-scale)
  - Measures of the home environment quality (FCI), time use survey
  - Mother's years of education, level of vocabulary ("Peabody") ...
- There are two main issues to deal with:
  - Multiple measures are likely to proxy common underlying constructs
  - Measures are imperfect proxies for skills and investments
- We tackle these issues using the latent factor approach of Cunha et al. (2010)

# Measurement equations

- For each $k = \{c, n\}$ and period $t$, we have measures of child's skill $\theta_{k,t}$:

$$m_{k,t,j} = \alpha_{k,t,j}\theta_{k,t} + \epsilon_{k,t,j}$$

  - $m_{k,t,j}$ is $j^{th}$ measurement of skill $\theta_{k,t}$
  - $\alpha_{k,t,j}$ is a factor loading
  - $\epsilon_{k,t,j}$ is the measurement error contained in $m_{k,t,j}$

- Similarly, for parental skills and investments:

$$\begin{aligned} \text{Parent's skill:} \quad m^P_{k,j} &= \alpha^P_{k,j} P_k + \epsilon^P_{k,j} \qquad (k = c, n) \\ \text{Investment:} \quad m^I_{t,j} &= \alpha^I_{t,j} I_t + \epsilon^I_{t,j} \end{aligned}$$

- The latent factors $\theta$ are the error-free measures of skills and investments we want to recover

# Identification of the measurement system

- The identification draws from the Kotlarsky theorem:
  - With two independent measurements per factor, the distribution of the unobserved factor and the measurement error can be identified non-parametrically up to a change of location

- We make some normalizations:
  - To set the factor scale: $\alpha_{k,t,1} = \alpha^I_{k,t,1} = \alpha^P_{k,t,1} = 1$
  - To set the factor location: $E(\theta_{k,t}) = E(\theta^I_{k,t}) = E(\theta^P_{k,t}) = 0$

- The approach can be generalized to allow for correlated measurement
  - Today's results assume uncorrelated measurement error across measures
  - The data collection design provides a unique opportunity to relax this assumption

# Estimation of the measurement system

- Recall, for example, the measurement equation for child's skill

$$m_{k,t,j} = \alpha_{k,t,j} \ln(\theta_{k,t}) + \epsilon_{k,t,j}$$

- We allow control and treated groups to have different factor distributions
  - For each group, we allow $\ln(\theta)$ to be distributed as a mixture of 2 normals

$$p_s(\ln \theta) = \tau_s \phi(\ln(\theta); \mu_{A,s}, \Sigma_{A,s}) + (1 - \tau_s)\phi(\ln(\theta); \mu_{B,a}, \Sigma_{B,s})$$

  - We normalize the factor means to 0 among one group (w.l.g., controls)

$$E_0(\ln(\theta)) = \tau_0 \mu_{A,0} + (1 - \tau_0)\mu_{B,0} = 0$$

- We assume that measurement error terms:
  - Follow a joint normal distribution (not necessary)
  - Are uncorrelated across measures (for now)
  - Are independent from factors (necessary)

# Estimation of the measurement system

- Under these assumptions, the measurements follow a mixture of normals
- This leads to the following likelihood function:

$$\begin{aligned} L_{i,s}(m) &= \int_\theta f(m_{i,s}|\theta)[\tau_s \phi_{A,s}(\theta_s) + (1-\tau_s)\phi_{B,s}(\theta_s)]d\theta = \\ &= \int_\theta [\tau_s f(m_{i,s}|\theta)\phi_{A,s}(\theta_s)d\theta + \int_\theta [\tau_s f(m_{i,s}|\theta)\phi_{B,s}(\theta_s)d\theta \end{aligned}$$

  where $f(m_{i,s}|\theta)$ is the density of the measurement system

- Each component is just a normal distribution with an analytical expression

# Table of Contents

# Controlling for endogenous investments

- Parents make investments in reaction to omitted inputs and unobserved shocks
- We need to control for the correlation between shocks and investments to identify the production functions
- We think of investments as being a function of:
  - All the factors that enter the production function
  - Whether the child was treated (received home visits)
  - Variables that shift investments but do not enter the production function
    - Family resources (wealth, income)
    - Family composition (child's birth order, mother's marital status)
    - Environmental variables (prices, distance to day care center)
- Resources are potentially endogenous, so we use resources measured at baseline

# Econometric approach to endogeneity

- We use a control function approach to account for the correlation between shocks and investments

- We specify a log-linear investment function:

$$\begin{aligned} ln(I_t) &= \lambda_{s,0} + \lambda_{s,1} ln(\theta_{c,t}) + \lambda_{s,2} ln(\theta_{n,t}) + \lambda_{s,3} ln(P_c) + \lambda_{s,4} ln(P_n) + \\ &+ \lambda_{s,5} ln(Z_t) + \eta_t \end{aligned}$$

  - $Z_t$ are instruments (village-level female and male wages, food prices, number of siblings, and a single child family indicator)

- We include the residuals of the investment equation, $\hat{\eta}_t$, as an additional regressor in the production function:

$$\begin{aligned} ln(\theta_{k,t+1}) &= ln(A_s) + \frac{1}{\rho_k} ln[\gamma_{k,1} \theta_{c,t}^{\rho_k} + \gamma_{k,2} \theta_{n,t}^{\rho_k} + \gamma_{k,3} I_{t+1}^{\rho_k} + \gamma_{k,4} P_c^{\rho_k} + \gamma_{k,5} P_n^{\rho_k}] \\ &+ \delta \hat{\eta}_t + v_{k,t} \end{aligned}$$

# Outline

# Data: Measures of skills and investments

| Factor | Survey | Measures |
|---|---|---|
| $\theta_{c,t+1}$ | FU | Bayley (cog, receptive lang, expressive lang, fine motor, gross motor) |
| | | MacArthur-Bates (words and phrases) |
| $\theta_{c,t}$ | BA | Bayley (cog, receptive lang, expressive lang, fine motor, gross motor) |
| | | MacArthur-Bates (words and understanding) |
| $\theta_{n,t+1}$ | FU | Bates (unsociable, difficult, unadaptable, unstoppable) |
| | | Rothbart (attention, sociable, inhibited) |
| $\theta_{n,t}$ | BA | Bates (unsociable, difficult, unadaptable, unstoppable) |
| $\theta_t^I$ | FU | FCI play materials (toys, books, drawing books...) |
| | | FCI play activities (reading, singing, naming...) |
| | | Detailed time diary of various activities w/ kid |
| $\theta_c^P$ | FU | Mother's vocabulary ("Peabody" scale) |
| | BA | Mother's completed years of education |
| | BA | Number of books for adults in the home (FCI) |
| | BA | Number of newspapers and magazines in the home (FCI) |
| $\theta_n^P$ | BA | CEDS Depression scale (10 questions) |

FU = Follow-Up; BA = Baseline

# Estimates of the joint distribution of latent factors

Factor means for the treated group relative to the control group
(normalized by the control group's standard deviation)

| Cognitive Skill (Follow-up) | Cognitive Skill (Baseline) | Non-cognitive Skill (Follow-up) | Non-cognitive Skill (Baseline) | Investment (Follow-up) | Mother's cognitive skill | Mother's non-cognitive skill |
|---|---|---|---|---|---|---|
| 0.083*** | 0.018 | 0.041*** | 0.018 | 0.260*** | -0.010 | 0.019 |
| (0.013) | (0.004) | (0.007) | (0.004) | (0.035) | (0.007) | (0.005) |

# The empirical importance of measurement error

- Recall the measurement equation

$$m_j = \alpha_j \ln(\theta) + \epsilon_j$$

- The variance of $m_j$ can be decomposed as:

$$Var(m_j) = \underbrace{\alpha_j^2 Var(\ln \theta)}_{\text{Signal}} + \underbrace{Var(\epsilon_j)}_{\text{Noise}}$$

- We can calculate the fractions of $Var(m_j)$ due to signal and noise:

$$\text{Signal:} \quad s^{\theta} = \frac{\alpha_j^2 Var(\ln \theta)}{\alpha_j^2 Var(\ln \theta) + Var(\epsilon_j)}$$

$$\text{Noise:} \quad s^{\epsilon} = \frac{Var(\epsilon_j)}{\alpha_j^2 Var(\ln \theta) + Var(\epsilon_j)}$$

- In the next tables, we report the fraction of the variance due to signal for the control group (very similar for treated)

# Percentage of Total Variance in Measurements due to Signal

**Measures of Cognitive and Non-Cognitive Skills at Follow-up**

| | Signal | Noise |
|---|---|---|
| Mac Arthur Words | 0.43 | 0.57 |
| Mac Arthur Phrases | 0.29 | 0.71 |
| Bayley Cognitive | 0.81 | 0.19 |
| Bayley Expressive Language | 0.81 | 0.19 |
| Bayley Receptive Language | 0.73 | 0.27 |
| Bayley Fine Motor | 0.62 | 0.38 |
| Bayley Gross Motor | 0.56 | 0.44 |
| Bates Unsociable | 0.15 | 0.85 |
| Bates Difficult | 0.58 | 0.42 |
| Bates Unadaptable | 0.09 | 0.91 |
| Bates Unstoppable | 0.43 | 0.57 |
| Rothbart Attention | 0.13 | 0.87 |
| Rothbart Inhibited | 0.55 | 0.45 |
| Rothbart Social | 0.04 | 0.96 |

# Percentage of Total Variance in Measurements due to Signal

**Measures of Cognitive and Non-Cognitive Skills at Baseline**

| | Signal | Noise |
|---|---|---|
| Mac Arthur Words | 0.24 | 0.76 |
| Mac Arthur Understand | 0.01 | 0.99 |
| Bayley Cognitive | 0.65 | 0.35 |
| Bayley Expressive Language | 0.62 | 0.38 |
| Bayley Receptive Language | 0.69 | 0.31 |
| Bayley Fine Motor | 0.56 | 0.44 |
| Bayley Gross Motor | 0.56 | 0.44 |
| Bates Unsociable | 0.08 | 0.92 |
| Bates Difficult | 0.49 | 0.51 |
| Bates Unadaptable | 0.19 | 0.81 |
| Bates Unstoppable | 0.21 | 0.79 |

# Percentage of Total Variance in Measurements due to Signal

**Measures of Investments**

| | Signal | Noise |
|---|---|---|
| Number of different play activities (FCI) | 0.85 | 0.15 |
| Number of different play materials (FCT) | 0.79 | 0.21 |
| Times read to kid in last 3 days | 0.62 | 0.38 |
| Times told a story in last 3 days | 0.62 | 0.38 |
| Times took kid outside in last 3 days | 0.37 | 0.63 |
| Times played with toys in last 3 days | 0.59 | 0.41 |
| Times named things to kids in last 3 days | 0.59 | 0.41 |
| Number of picture books | 0.43 | 0.57 |
| Number of books to paint and draw | 0.45 | 0.55 |
| Number of toys to move | 0.47 | 0.53 |
| Number of toys to learn shapes | 0.60 | 0.40 |

# Percentage of Total Variance in Measurements due to Signal

**Measures of Mother's Cognitive and Non-Cognitive Skills**

| | Signal | Noise |
|---|---|---|
| Mother's "peabody score" | 0.61 | 0.39 |
| Mother's years of education | 0.62 | 0.38 |
| Number of adult books at home | 0.42 | 0.58 |
| Number of magazines and newspapers at home | 0.22 | 0.78 |
| CESD - A | 0.18 | 0.82 |
| CESD - B | 0.24 | 0.76 |
| CESD - C | 0.00 | 1.00 |
| CESD - D | 0.42 | 0.58 |
| CESD - E | 0.21 | 0.79 |
| CESD - F | 0.28 | 0.72 |
| CESD - G | 0.20 | 0.80 |
| CESD - H | 0.15 | 0.85 |
| CESD - I | 0.23 | 0.77 |
| CESD - J | 0.26 | 0.74 |

# Outline

# The investment equation

| | Point estimate | 90% CI |
|---|---|---|
| Intercept | 0.000 | [-0.424 ; 0.483] |
| Treatment (home visits) | 0.261 | [0.152 ; 0.358] |
| Cognitive skill (t) | 0.041 | [-0.112 ; 0.162] |
| Non-cognitive skill (t) | -0.212 | [-0.391 ; 0.131] |
| Mother's cognitive skill | 0.614 | [0.453 ; 0.732] |
| Mother's non-cognitive skill | 0.085 | [-0.086 ; 0.192] |
| Average female wages in village | -1.452 | [-4.548 ; -0.691] |
| Average male wages in village | 1.485 | [0.723 ; 4.528] |
| Average food price in village | -0.050 | [-0.091 ; 0.023] |
| Number of siblings | 0.022 | [-0.032 ; 0.093] |
| Single child family | 0.074 | [0.037 ; 0.100] |

# CES production function for cognitive skills (t+1)

| | No control function | W/ Control Function |
|---|---|---|
| TFP parameter | **1** | **0.999** |
| | [0.988,1.006] | [0.986,1.006] |
| TFP parameter x Treat | **0.04** | **0.014** |
| | [-0.008,0.084] | [-0.049,0.074] |
| Cognitive skill (t) | **0.811** | **0.801** |
| | [0.667,0.934] | [0.65,0.933] |
| Non-cognitive skill (t) | **0.028** | **0.035** |
| | [-0.16,0.162] | [-0.162,0.169] |
| Mother's cognitive skill (t) | **0.063** | **-0.004** |
| | [0.002,0.176] | [-0.144,0.149] |
| Mother's non-cognitive skill (t) | **0.024** | **0.007** |
| | [-0.063,0.148] | [-0.099,0.132] |
| Investment (t) | **0.103** | **0.204** |
| | [0.047,0.162] | [0.053,0.424] |
| Number of siblings (t) | **-0.049** | **0.016** |
| | [-0.069,-0.01] | [-0.022,0.052] |
| Single child family (t) | **0.021** | **-0.059** |
| | [-0.015,0.056] | [-0.087,-0.016] |
| Control function | **-** | **-0.109** |
| | | [-0.343,0.06] |
| Complementarity parameter | **0.091** | **0.109** |
| | [-0.096,0.327] | [-0.056,0.268] |
| Elasticity of substitution | **1.1** | **1.122** |
| | [0.912,1.485] | [0.947,1.366] |

# CES production function for non-cognitive skills (t+1)

| | No control function | W/ Control Function |
|---|---|---|
| TFP parameter | **1.013** | **1.006** |
| | [1,1.027] | [0.996,1.016] |
| TFP parameter x Treat | **-0.009** | **-0.066** |
| | [-0.034,0.022] | [-0.143,-0.006] |
| Cognitive skill (t) | **0.165** | **0.141** |
| | [0.102,0.297] | [0.066,0.286] |
| Non-cognitive skill (t) | **0.691** | **0.703** |
| | [0.439,0.716] | [0.419,0.717] |
| Mother's cognitive skill (t) | **-0.05** | **-0.198** |
| | [-0.077,0.037] | [-0.346,-0.035] |
| Mother's non-cognitive skill (t) | **0.063** | **0.026** |
| | [0.004,0.213] | [-0.042,0.178] |
| Investment (t) | **0.121** | **0.349** |
| | [0.06,0.186] | [0.164,0.654] |
| Number of siblings (t) | **-0.008** | **0.008** |
| | [-0.028,0.026] | [-0.054,0.047] |
| Single child family (t) | **0.017** | **-0.029** |
| | [-0.034,0.056] | [-0.057,0.006] |
| Control function | **-** | **-0.241** |
| | | [-0.591,-0.068] |
| Complementarity parameter | **-0.276** | **-0.097** |
| | [-0.567,-0.009] | [-0.244,0.018] |
| Elasticity of substitution | **0.784** | **0.912** |
| | [0.638,0.991] | [0.804,1.018] |

# Cobb-Douglas production function for cognitive skills (t+1)

| | No control function | W/ Control Function |
|---|---|---|
| TFP parameter | **1.003** | **1.003** |
| | [0.994,1.006] | [0.993,1.007] |
| TFP parameter x Treat | **0.04** | **0.015** |
| | [-0.008,0.084] | [-0.058,0.07] |
| Cognitive skill (t) | **0.809** | **0.799** |
| | [0.666,0.93] | [0.64,0.929] |
| Non-cognitive skill (t) | **0.027** | **0.032** |
| | [-0.158,0.16] | [-0.167,0.169] |
| Mother's cognitive skill (t) | **0.062** | **0** |
| | [0.002,0.176] | [-0.151,0.139] |
| Mother's non-cognitive skill (t) | **0.026** | **0.01** |
| | [-0.061,0.145] | [-0.087,0.13] |
| Investment (t) | **0.105** | **0.2** |
| | [0.048,0.165] | [0.048,0.446] |
| Number of siblings (t) | **0.021** | **0.017** |
| | [-0.015,0.056] | [-0.023,0.051] |
| Single child family (t) | **-0.05** | **-0.059** |
| | [-0.071,-0.011] | [-0.088,-0.019] |
| Control function | **-** | **-0.101** |
| | | [-0.341,0.047] |
| Elasticity of substitution | **1** | **1** |

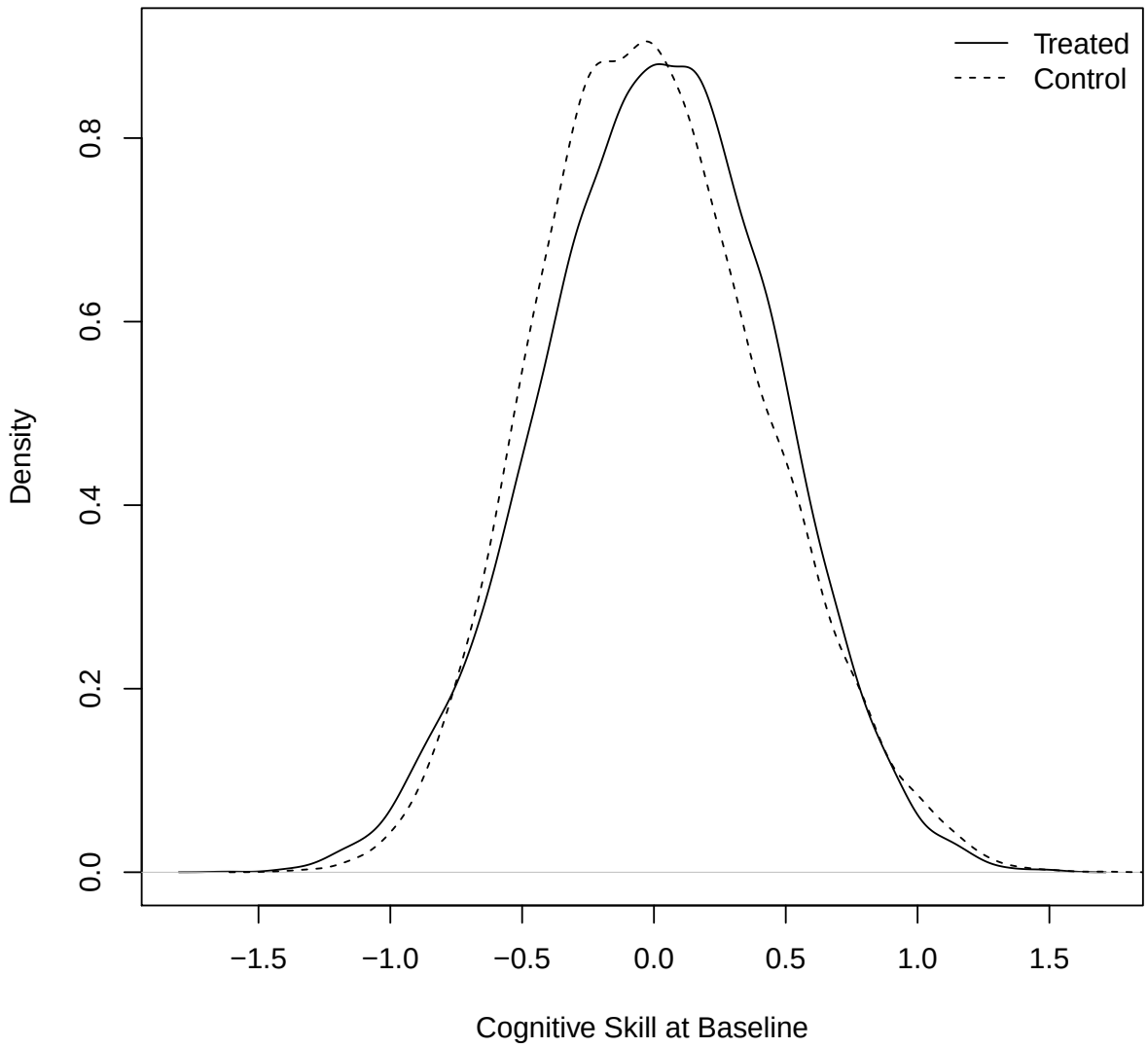
# Cobb-Douglas production function for non-cognitive skills (t+1)

| | No control function | W/ Control Function |
|---|---|---|
| TFP parameter | **1.001** | **1.001** |
| | [0.995,1.004] | [0.993,1.006] |
| TFP parameter x Treat | **-0.008** | **-0.065** |
| | [-0.032,0.023] | [-0.149,-0.006] |
| Cognitive skill (t) | **0.167** | **0.143** |
| | [0.101,0.299] | [0.067,0.287] |
| Non-cognitive skill (t) | **0.696** | **0.709** |
| | [0.443,0.724] | [0.424,0.723] |
| Mother's cognitive skill (t) | **-0.051** | **-0.202** |
| | [-0.077,0.038] | [-0.36,-0.03] |
| Mother's non-cognitive skill (t) | **0.061** | **0.023** |
| | [0.004,0.214] | [-0.043,0.168] |
| Investment (t) | **0.116** | **0.349** |
| | [0.057,0.18] | [0.174,0.659] |
| Number of siblings (t) | **0.019** | **0.009** |
| | [-0.035,0.057] | [-0.055,0.05] |
| Single child family (t) | **-0.009** | **-0.031** |
| | [-0.03,0.028] | [-0.058,0.008] |
| Control function | **-** | **-0.245** |
| | | [-0.594,-0.074] |
| Elasticity of substitution | **1** | **1** |

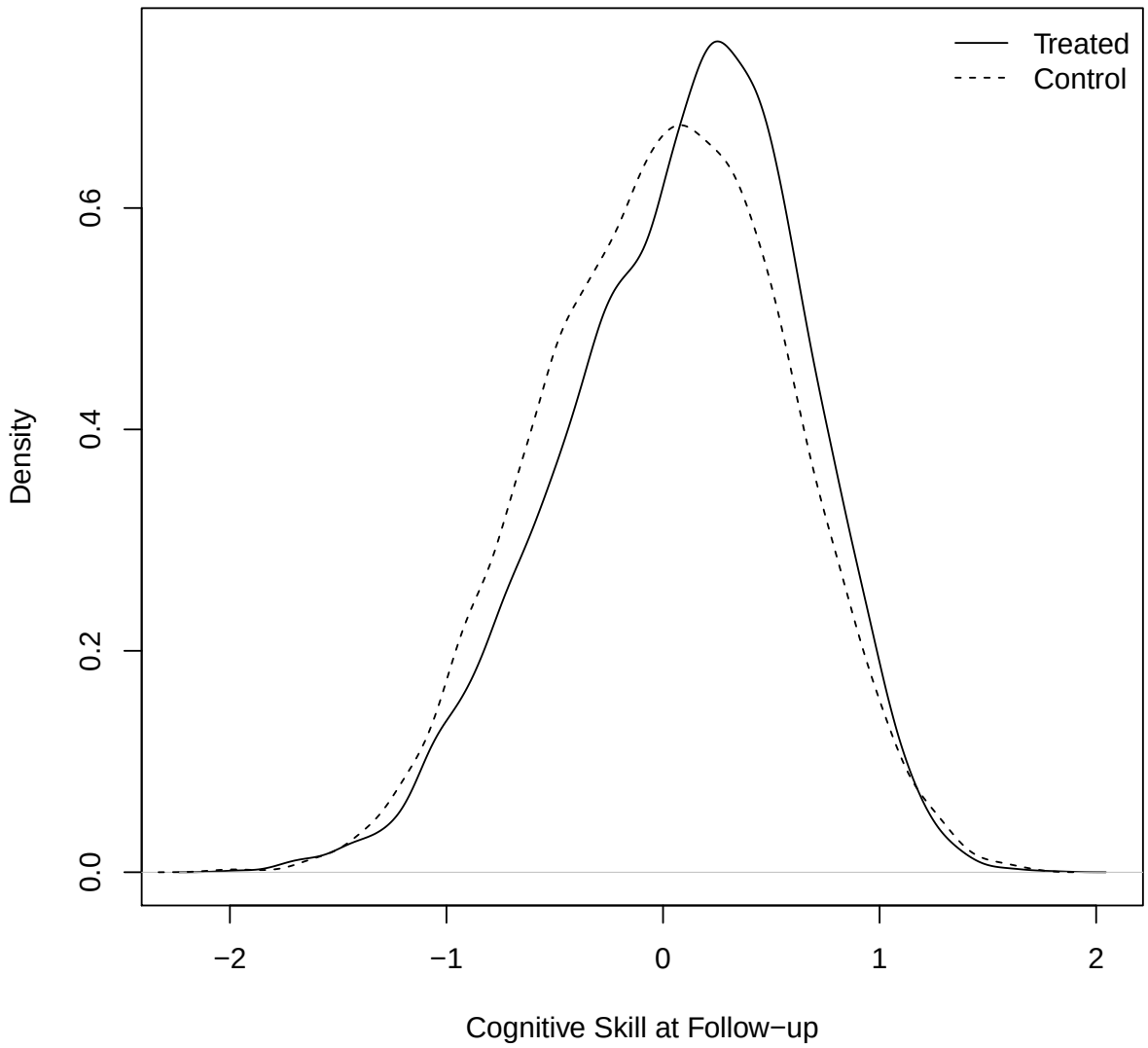
# Using the framework to understand the intervention

- In the model, the intervention could have operated through two channels:
  1. By shifting the distribution of parental investments
  2. By boosting TFP parameter
- The model estimates suggest that it was mostly through channel 1
- By simulating the model, we can assess how well the model does explains the effect of the intervention

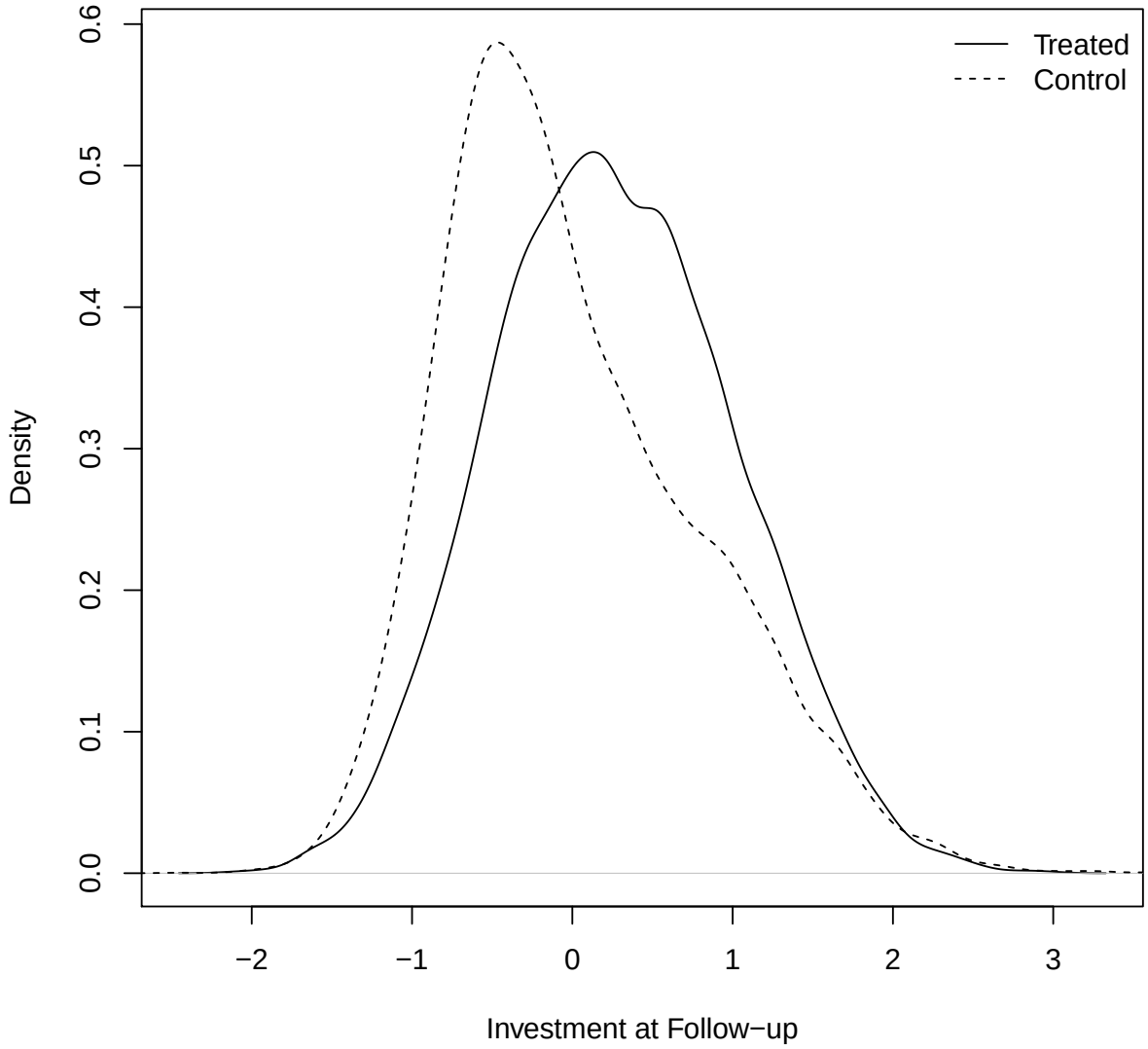
# Distribution of cognitive skills at baseline



# Distribution of cognitive skills at follow up



# Distribution of the investment factor



# Using the model to predict impacts of the intervention

Difference in average cognitive and non-cognitive skills

between treated and controls in the data and as predicted by the model

| | Cognitive | non-cognitive |
|---|---|---|
| Data | 0.084 | 0.029 |
| | (0.010) | (0.006) |
| Model | 0.082 | 0.022 |
| | (0.076) | (0.050) |

# Outline

# Summary

- We estimate production functions for the development of cognitive and non-cognitive skills
- We extend the framework of Cunha et al. (2010) to understand the role an early childhood intervention in Colombia has played in the formation of human capital
- We find evidence that:
  - Strong self-productivity effects
  - Cross-productivity effects in the production of non-cognitive skills
  - Parental investments matter and are complementarity with other inputs
  - Investments matter and parents seem to react to shocks to mitigate their impacts
- The impact of the intervention can be explained by a shift in the distribution of investments

# Next steps

- Relax assumptions underlying the estimation of the production function
  - Account for correlated measurement error
  - Explore several dimensions of investments (time vs. material, mother's time vs. other's time)

- Exploit future follow-up data in order to further investigate the dynamics
  - Are the impacts sustained over time?
  - How does the technology change with age?
  - To what extent are investments complementary over time?