NBER WORKING PAPER SERIES

WINDOWS OF OPPORTUNITY IN EARLY CHILDHOOD: A DYNAMIC MODEL OF MULTIDIMENSIONAL DEVELOPMENT

Orazio Attanasio
Raquel Bernal
Michele Giannola
Milagros Nores

Working Paper 35818
http://www.nber.org/papers/w35818

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2026

We are grateful to many colleagues for a number of conversations. In particular, we would like to thank Sarah Cattan, Flavio Cunha, Jim Heckman, Costas Meghir and Emily Nix for useful discussions on measuring child development at different ages. We also received very valuable comments from Lance Lochner and 4 anonymous referees. We used Claude to improve the style of some paragraphs in the text. We would also like to thank AEIOTU's staff for the support. Attanasio's research was partially funded by ERC Advanced Grant AdG – 695300. Nores and Bernal's research underlying this work was supported by the Jacobs Foundation (Grant No. 209-805-1), the UBS Optimus Foundation, the LEGO Foundation, and the Inter-American Development Bank (PO 209-805-1). The views expressed are those of the authors and do not necessarily represent those of the funders. The study obtained ethical approvals from Universidad de Los Andes, Comite de Etica (IRB Approval Number 032-2009) and Rutgers IRB (IRB Approval Number E09-568). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Orazio Attanasio, Raquel Bernal, Michele Giannola, and Milagros Nores. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Windows of Opportunity in Early Childhood: A Dynamic Model of Multidimensional Development
Orazio Attanasio, Raquel Bernal, Michele Giannola, and Milagros Nores
NBER Working Paper No. 35818
September 2026
JEL No. I15, I25, I32, J13, J24, O15

ABSTRACT

Using child development data collected during the evaluation of an early childhood education program, this paper estimates a dynamic model of human development from ages 1 to 6. We model three latent factors—health, cognitive skills, and socio-emotional skills—and allow them to interact with each other and with multiple inputs to shape subsequent development. The richness and high-frequency nature of the data permit a flexible specification that (i) allows for complex dynamics, (ii) lets each factor con- tribute to the production of the others, (iii) permits age-varying production technologies, and (iv) treats parental investment as endogenous. We use the estimated model to identify "windows of opportunity" in which parental interventions, or programs that increase investment, are especially effective. Results indicate that developmental dynamics are richer than typically assumed, with important implications for persistence: the impact of inputs can decay or endure in ways that vary across domains and over age. Accounting for endogenous investment is consequential for inference.

Orazio Attanasio
Yale University
Department of Economics
and CEPR
and also NBER
orazio.attanasio@yale.edu

Michele Giannola
University of Naples Federico II
Department of Economics and Statistics
and Institute for Fiscal Studies and
CSEF
michele.giannola@unina.it

Raquel Bernal
Universidad de los Andes
and CEDE
rbernal@uniandes.edu.co

Milagros Nores
NIEER, Rutgers University
mnores@nieer.org

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/1903

1 Introduction

How different domains of early human capital—cognitive skills, socioemotional development, and health—respond to parental investment, and how this responsiveness evolves as children grow, are questions of first-order importance for understanding skill formation and designing effective policies. Recent contributions have emphasized the existence of windows of opportunity during early childhood—periods in which investments may have particularly large and lasting effects on development. The timing and nature of these windows depend critically on the features of the developmental process: how the productivity of parental investments changes with age, and how the persistence of each dimension of development evolves over the early years of life.

Persistence, that is, the extent to which current levels of development shape future development, is particularly important. Low persistence implies that development remains relatively malleable, allowing subsequent investments to modify children's developmental trajectories. As persistence increases, the developmental process becomes increasingly self-reinforcing: the effects of early investments become more enduring, but so too do the consequences of early disadvantage, making them progressively more difficult to offset. The timing of windows of opportunity therefore depends on the interaction between the age profile of investment productivity and the evolution of persistence over the life cycle.

This paper provides new evidence on these questions. We exploit a rich longitudinal dataset that follows children *annually* from age 1 to age 7, with multiple measures of cognitive development, socioemotional skills, and health at each wave, as well as parental contributions. This combination — high frequency, long time horizon, and simultaneous measurement of three domains — is rare in this literature. It allows us to estimate production functions for child development that are considerably richer than those previously analyzed: tracing how the persistence of each dimension evolves over early childhood, documenting how the productivity of parental investment changes with age, and capturing potential cross-dimensional interactions.

Our main findings reveal strikingly different dynamics across the three dimensions. For **cognitive development**, parental investment is highly productive in the early years — ages 2 through 4 — but its effectiveness declines sharply thereafter. At the same time, persistence is relatively low at early ages and jumps substantially at age 5, suggesting the process becomes increasingly self-reinforcing as children approach school age. Taken together, these two patterns define a narrow and early window of opportunity for cognitive development. **Socioemotional skills** present a different picture: parental investment remains productive across all ages, while persistence is comparatively low and stable throughout — implying a

much longer window during which the process remains malleable and responsive. **Health**, by contrast, responds to parental investment only at age 2, but becomes highly persistent from age 3 onward, approaching a unit root. We find little evidence of cross-domain interactions. Together, these patterns identify a clear hierarchy of windows of opportunity that differs sharply across domains — with direct implications for the design and timing of early childhood interventions.

These findings emerge from a framework that addresses three challenges that have limited progress in this literature. First, we address the **measurement challenge** posed by the fact that available tests change with the child's age: making developmental outcomes comparable across ages requires careful handling of the identifying restrictions of the measurement system. We propose an approach that constructs comparable metrics for each domain across ages with a minimal set of assumptions — a prerequisite for interpreting persistence parameters, and an improvement on existing research showing that strong normalization assumptions can generate significant biases (Agostinelli and Wiswall, 2016, 2025). Second, rather than focusing on a single domain of development, we **estimate all three domains jointly**, allowing for potential cross-domain spillovers while remaining agnostic about whether they exist. Third, we account for the **endogeneity of parental investment**, instrumenting it with household financial resources at each age.

A growing literature developed in the last few decades has laid the foundations for our work. A substantial body of evidence in economics and developmental psychology has stressed the importance of the early years for parental investment and the malleability of child development (Cunha and Heckman, 2007; Heckman, 2006; Nelson, 2000; Shonkoff and Phillips, eds, 2000). Our contribution, however, analyzes the early years at a much finer level of temporal disaggregation and, to the best of our knowledge, is the first to exploit evidence at this high frequency to identify what the literature has labeled as *windows of opportunity*.

From an econometric point of view, Cunha et al. (2010) established a non-parametrically identified model of multidimensional child development accounting for measurement error and flexible functional forms. Attanasio et al. (2020a) extended this framework with a more flexible estimation approach that pays careful attention to investment endogeneity, with applications to India, Peru, and Ethiopia (Attanasio et al., 2017). Attanasio et al. (2020b) applied these methods to interpret the impacts of an early childhood intervention in Colombia, conducting a formal mediation analysis in the spirit of Heckman et al. (2013). Our paper builds on and extends this line of work, exploiting a dataset with substantially richer dynamics than those previously used and proposing a normalization strategy that makes dynamic comparisons across ages consistent and easier to interpret.

After showing that our theoretical model fits the available data well, we use the esti-

mates of the process of child development, in its three domains, to perform some simulation exercises. First, we compute the marginal productivity of parental investment for the three domains of human development considered and document how they change over time. Finally, we compute impulse response functions providing strong support for the proposition that policies that affect parental investment (or more generally child development) are most effective and persistent between the ages of 2 and 4.

The remainder of the paper is organized as follows. Section 2 presents the conceptual framework. Section 3 discusses the measurement approach and normalization. Section 4 describes the data. Section 6 presents estimates of the investment and production functions. Section 7 reports counterfactual experiments varying the timing and duration of parental investment. Section 8 concludes.

2 The production function of human development

In this section, we present a framework for modeling child development: a production function mapping key inputs into developmental outcomes. We first outline the domains of interest, then specify the production function and examine how its parameters may vary across developmental stages. Special attention is given to parental investment, which — as a product of individual choice — is likely correlated with unobservable factors; we address this endogeneity concern using a control function approach. We conclude with a discussion of how the dynamics of child development are reflected in the estimated evolution of the model's parameters.

2.1 Different domains: cognition, socioemotional skills and health

As mentioned in the introduction, human development is widely recognized as a multidimensional process that encompasses multiple skills and traits. Recognizing these multiple domains is important both when studying the process of development itself and because of the role they play in determining different adult outcomes and individual wellbeing.

We model the dynamic evolution of three core domains: cognitive skills, socio-emotional skills, and health. This choice is partly dictated by the availability of suitable measures in our data. Richer datasets could support a more detailed characterization of the developmental process—for example, by distinguishing cognition from language development or, more importantly, by modeling internalizing and externalizing socio-emotional skills separately, as in Attanasio, Blundell, Conti and Mason (2018). However, our specification is among the richest in the existing literature, particularly in studies based on data from developing

countries.[1]

A unique feature of our data is the frequency with which measures of child development are collected. Children are observed annually from age 6 months to age 7, covering a wide spectrum of ages. Although the panel is unbalanced — many children are followed for approximately three years, but some for substantially longer periods[2] — the sample is large and the coverage broad. This high frequency is what allows studying in detail how the process of child development changes over time during the first 7 years of life, and to identify the windows of opportunity for intervention that we referred to in the introduction.

2.2 Specifying the production function

Section 3 maps the available measures to the three developmental domains — or latent factors — that we model. This structure allows us to examine how these factors interact with one another and with other variables across the developmental process, and how those relationships evolve with age. Such a framework is valuable both academically and from a policy standpoint, as it helps identify which skills matter most at different stages of development — knowledge that can inform the design of targeted and/or or domain-specific interventions. The estimated process of human development is represented by the following equation:

$$\boldsymbol{\theta_{a+1}} = f_a(\boldsymbol{\theta_a}, \boldsymbol{X_a}, \boldsymbol{Z_a}, \boldsymbol{\varepsilon_{a+1}}) \tag{1}$$

where a denotes age, $\boldsymbol{\theta_a}$ is a vector of latent factors representing the different dimensions of child development, and $\boldsymbol{\varepsilon_{a+1}}$ captures unobservable shocks to the developmental process. The other inputs in equation (1) are partitioned into two vectors: $\boldsymbol{X_a}$, which contains variables chosen by agents — most importantly, parental investment[3] — and $\boldsymbol{Z_a}$, which contains environmental and background factors taken, in the present context, as given, such as parental education, family composition and, importantly, the assignment to the *aeioTU* program. This distinction matters because $\boldsymbol{X_a}$ may be correlated with $\boldsymbol{\varepsilon_{a+1}}$ — that is, parental choices may respond to unobserved shocks to child development — making these variables *endogenous*.

The vector $\boldsymbol{Z_a}$ includes environmental and background factors such as family composition and parental education. The key endogenous variable in $\boldsymbol{X_a}$ is parental investment.

[1] Attanasio et al. (2020a) and Attanasio et al. (2017), for example, use Young Lives data to jointly model cognitive skills and health. Cunha et al. (2010) use NLSY data to study cognition and socio-emotional skills at substantially later ages than those considered here. Rubio-Codina, Araujo, Attanasio, Munoz and McGregor (2017) discuss measurement challenges in assessing early childhood development.

[2] These are the youngest children at the beginning of the study.

[3] Empirically, in this paper, we consider as parental investment toys, children books and the like. A complete list of the items that we use for the purpose is in Table 2

While parental behavior is widely recognized as an important determinant of child development, investment decisions are likely to respond to the child's developmental trajectory — parents may invest more when they observe their child falling behind, or less when development appears on track. This dependence between investment choices and the unobserved (to the researcher) shocks in $\boldsymbol{\varepsilon_{a+1}}$ makes identifying the causal effect of parental investment on child development particularly challenging.

At this point, we keep the function f_a, which determines the evolution of human development over time, purposely vague, with the idea of wanting to use a flexible specification where different inputs may play an important role. We note the subscript a to the function f to stress that we let this function to change with age. Indeed, one of main goals of this study is to establish how the function f changes with age.

Estimating the function f_a in equation (1) poses three challenges. First, while a fully non-parametric approach is conceivable in principle, the dimensionality of the problem makes it infeasible in practice. Therefore, we adopt parametric specifications that impose enough structure to make the production function estimable while preserving a meaningful degree of flexibility. Second, as mentioned above, the variables in $\boldsymbol{X_a}$ might be chosen in response to a child's developmental progress and are therefore endogenous, complicating the identification of their causal role. Third, some of $\boldsymbol{\theta_a}$, $\boldsymbol{X_a}$, or $\boldsymbol{Z_a}$ is directly observable; instead, we observe more or less noisy *markers* of these latent variables. We address the first two challenges in the remainder of this section and turn to the measurement issues in Section 3.

2.2.1 Functional form assumptions

Given the data available, we model three factors, which we denote as 'cognitive skills' (labeled C), 'socio-emotional skills' (labeled S) and 'health' (labeled H). For each of these factors and for each age considered, we estimate different functional forms for the production function. The first specification we consider for each of the three human capital domains considered is the Constant Elasticity of Substitution (CES) specification.

A CES specification can accommodate a large number of inputs whose effect on the total outcome are neither additive nor separable. One can impose the assumption of constant return to scale, but this restriction is not necessary. The CES allows for interactions among different inputs. However, a single parameter governs the degree of substitutability (or complementarity) among all the inputs entering the function. A possible generalization, at the cost of introducing important non-linearities, is the nested CES, which has been used, among others, by Attanasio et al. (2017). In this case, depending on the number of nests, several parameters determine the substitutability among different inputs. The Cobb-Douglas specification arises as a special case of the CES when the elasticity of substitution between

inputs equals one.. In this case, the specification is log-linear in the parameters of interest.

Despite the flexibility of the CES specification and the simplicity of the Cobb-Douglas function, these functional forms impose strong restrictions on the substitutability of different inputs. For this reason, we use a Transcendental Logarithmic production function (translog for short) that allows us capturing different degrees of substitutability between inputs flexibly (see, e.g., Christensen et al. (1973)). The translog specification (with some specificity in notation) is given by:

$$\begin{aligned} \ln \theta_{k,a+1} = \quad & \gamma_a^{k,C} \ln \theta_{C,a} + \gamma_a^{k,S} \ln \theta_{S,a} + \gamma_a^{k,H} \ln \theta_a^H + \gamma_a^{k,I} \ln \theta_{I,a} + \qquad k \in \{C,S,H\} \qquad (2) \\ & + \tfrac{1}{2} \textstyle\sum_q \sum_r \gamma_a^{k,q,r} \ln \theta_{q,a} \ln \theta_{r,a} + \delta_a^{k,0} + \delta_a^{k,Z\prime} \boldsymbol{Z_a} + \varepsilon_{a+1}^k \qquad q,r \in \{C,S,H\}. \end{aligned}$$

with $\gamma_a^{k,q,r} = \gamma_a^{k,r,q}$. $\theta_k, k = C, S, H$ are the 3 domains of development we consider, while θ_I is parental investment. The vector $\boldsymbol{Z_a}$ represents control variables including family composition variables, parental education and the assignment to the *aeioTU* program and other variables mentioned in the Tables. The log-linearity of this specification greatly simplifies the estimation of the key parameters. Furthermore, the translog function reduces to the Cobb-Douglas when all $\gamma_a^{k,q,r} = 0$.

This framework permits the estimation of age-specific complementarities and spillovers across skills - for example, the contribution of socio-emotional skills to subsequent cognitive development - as well as changes over time in the productivity of parental investment. Modeling these interactions explicitly is central to our analysis. In this respect, the translog specification provides substantially greater flexibility than the CES specification, in which all interactions are constrained to operate through a single parameter. As discussed above, the estimated levels and age profiles of these parameters have direct and important policy implications.

Our empirical analysis focuses on estimating the magnitude of the coefficients associated with the key inputs in the child development production function and, crucially, on how these coefficients evolve over time. In particular, we study three related features of the technology: (i) **the degree of persistence in each domain of human capital**, captured by the coefficients on lagged measures of the various domains of development, which are allowed to vary across ages; (ii) **the dynamic interactions across different domains of development**; and (iii) **the role of parental investment** at different stages of childhood. We discuss these issues next.

2.2.2 The dynamics of human development: persistence and non-stationarity

An important and, so far, relatively unexplored feature of the development process is how its dynamic properties change with age. Often, again due to the availability of data, researchers focus on models where the level of human development in various domains at age a depends on the level of human capital at age $a-k$ (and other variables), where $a-k$ is the age at which data are available. Such assumptions on the nature of persistence of the process are very strong and may be inconsistent with the evidence from certain interventions that show an effect in the short run that seems to disappear in the medium run and re-appear in the long run, probably indicating the presence of complex dynamic effects. Importantly, very few data sets allow studying how the persistence of the process and the effect of its different drivers change through various ages. Yet these key parameters are crucial for determining when windows of opportunities matter for child development and for policies aimed at improving child development for disadvantaged children.

As mentioned above, in the data we use, children are observed at a relatively high frequency and cover a wide range of ages, from 6 months to 7 years. These features allow us to study in detail how the process of child development changes, allowing us to identify the *windows of opportunities* for interventions we referred to earlier.

We also investigated whether the process of child development satisfies the Markov property, according to which, conditional on variables measured at age $a-1$ variables, variables measured at earlier ages do not predict development at age a. However, our results indicated that only for a few ages and some domains, the coefficients of development at $a-2$ were significantly different from zero. Furthermore, even when these additional lags are significant, their inclusion in the statistical model does not substantially change the identification of the windows of opportunity to stimulate parental investments considered in Section 7. Given this evidence, we focus on how the parameters of the process of child development change over time, and report the presence of longer lags and, therefore, a violation of the Markov property in the Appendix.[4]

2.2.3 The endogeneity of parental investment

Parents make choices that reflect their resources (both financial and time), preferences, and beliefs. One implication is that the correlation between parental investments and child outcomes does not necessarily have a causal interpretation. In other words, OLS or NLLS estimates of the coefficients γ in equation (2) may be biased. If so, understanding the

[4]In Section 7, we present goodness of fit evidence as well as the identification of *windows of opportunity* using both models with one and two lags of past developmental outcomes.

direction and magnitude of this bias, when estimable, can provide information on about what drives parental behavior. For example, a downward bias in the estimated marginal product of investment may indicate compensatory behavior, whereby parents increase investments in response to negative shocks on their children's development.

To obtain consistent estimates of the coefficients of the production function, it is useful to sketch a model of parental behavior. We assume parents maximize an objective function that depends on child development and their own consumption, subject to budget and time constraints and a (perceived) association between child development and parental investments. This optimization defines a 'parental investment function' that depends on resources, the relative price of parental investment, and other variables perceived to influence child development. The parameters of this function depend on parental preferences and beliefs about the process of child development.

A feasible approach to estimate the parameters of such a production function in this context would be to use an IV or, in the case of a non-linear function, a control function approach. For such an approach to work, it is necessary to identify variables that drive investment, but that can be plausibly excluded from the production function of child development. Once again, we note that, if the objective is to estimate the parameters of the 'true' production function, rather than the complete model, it is not necessary to specify a completely accurate investment function, assume that parents know the 'true' production function or estimate the complete model. Instead, one can estimate an 'investment' function that includes some variables that drive investment and are not direct determinants of a child's human capital.[5] This approach leads us to estimate the following approximation to the investment function:

$$\theta_{I,t} = \pi_0 + \pi_{C,t}\theta_{C,t} + \pi_{S,t}\theta_{S,t} + \pi_{H,t}\theta_{H,t} + \pi_{P,t}\theta_P + \pi'_{G,t}\boldsymbol{G_t} + \pi_{Z,t}\boldsymbol{W_t} + u_t \tag{3}$$

where the vector $\boldsymbol{G_t}$ are variables that affect the production function, while $\boldsymbol{W_t}$ are the instruments that do not enter the production function directly. Such an equation should be interpreted as a reasonable approximation (not necessarily consistent) to the investment function. Estimating it allows to obtain estimates of u_t, which can then be used to construct a control function to add to the production function equation.

In the context of a standard model of child development and parental investments, two variables that could be considered as 'plausible instruments' are the *price of investment* and the *financial resources* available to the household. As discussed in Attanasio et al. (2020b), prices could constitute a more attractive instrument than financial resources, as the former

[5]In linear models, the consistency of estimates so obtained can be easily proven. The issue is subtler in non-linear model and one might have to show the robustness of the results with simulations.

are plausibly taken as given by households, while the latter could proxy for some omitted input in the production function.

Unfortunately, in the context analyzed in this paper, there is likely limited variation in prices, as the children in our sample live in the same city across two neighborhoods. We therefore use variables related to financial resources as instruments and discuss the potential biases and interpretation of results in Section 6. Our estimates on the effectiveness of parental investment, therefore, should therefore be interpret with caution. It is worth noting, however, that Attanasio et al. (2020b), conduct an extensive analysis of instruments used to identify the effect of parental investments and obtain similar estimates of the relevant coefficient whether identification relies on prices or resources and that differ substantially from OLS estimates , which do not account for the endogeneity of parental investment.[6]

3 Latent factors and measurement

In this section, we first describe the measurement system that links the latent factors of interest to the available measures. We then discuss the normalization imposed on the system and how we address that, for several developmental domains, no single measure is available over the full age range covered by our data. Finally, we present the estimation of the child development production functions and discuss how we address the endogeneity of some inputs. The choice of the developmental domains we consider, as well as the corresponding measurement strategies, is discussed in detail in Section 4.2.

Several inputs to the production and investment functions, as well as their outputs, are not directly observed. Instead, the data provide multiple noisy measures that are informative about the underlying latent factors. We follow the approach developed in Cunha et al. (2010), Attanasio et al. (2020a), and Attanasio et al. (2020b), estimating a measurement system that aggregates the available information while explicitly accounting for measurement error. Estimation of this measurement system yields the joint distribution of the relevant latent factors, which is then used in the estimation of the production functions.

A key challenge in implementing this framework is the need to address normalization and location issues, particularly in a longitudinal setting. The problem is especially acute when no single measure is available over the entire age range and, even when such a measure exists, its relationship with the underlying latent factor is unlikely to remain stable as children age. Because our analysis focuses on the dynamic evolution of skills, imposing a consistent metric across ages is not merely a technical convenience but a prerequisite for meaningful

[6] Attanasio et al. (2020b) also discuss the possibility that the instruments they use are 'weak' and therefore introducing a bias in the estimates. They show that instrument weakness is not an issue in their context.

inference.

3.1 The measurement system

Following Cunha et al. (2010), we specify the measurement system in terms of the log of the latent factors, ensuring that the factors only take positive values, as required by the specification of the production function. Let $m^j_{k,t}$ be a continuous measure j at age t for factor k.[7] Assuming a log-linear relationship between the measures and the factors with additive measurement error, we can write:

$$m^j_{k,t} = \alpha^j_{k,t} + \lambda^j_{k,t} \ln \theta_{k,t} + \epsilon^j_{k,t} \tag{4}$$

where $\alpha^j_{k,t}$ is an intercept term and $\lambda^j_{k,t}$ is a factor loading. The intercept captures the *discriminating power* of measure $m^j_{k,t}$: a measure with a low value of $\alpha^j_{k,t}$ is one that, for a given level of the factor $\theta_{k,t}$, obtains a lower value. The factor loading captures the *saliency* of variable j for factor k: a measure with a high value of $\lambda^j_{k,t}$ is one that varies greatly with the variation in the factor. Both the intercept and the factor loading are indexed by t, reflecting the fact that these are allowed to vary by age, as to capture the potentially different informativeness of the same measure over time. The term $\epsilon^j_{k,t}$ represents the measurement error for the variable j and prevents a deterministic relationships between the latent factors and their relevant markers. This is assumed to be mean zero and independent of the latent factor. Furthermore, the error terms are assumed to be independent between each other ($\epsilon^j \perp\!\!\!\perp \epsilon^{j'}$) and over time ($\epsilon_t \perp\!\!\!\perp \epsilon_{t'}$). These assumptions can be somewhat relaxed, although identification of the model's parameters requires measurement errors to be independent across at least some measures.

An important feature of our approach is that, having estimated the parameters of equation (4), it is possible to use a subset of the measures considered in this system to obtain estimates of the relevant latent factors. Indeed, the model in equation (4) can be considered as an efficient way to aggregate the information conveyed by the available measures on the factors of interest, which can be different domains of child development, parental investment or other variables.

Under certain conditions on the number of available measures, it is possible to identify the parameters in equation (4), without imposing any assumptions on the distribution of the latent factor $\theta_{k,t}$. However, in practice, researchers often impose specific assumptions on

[7] If the available measure $m^j_{k,t}$ is discrete (binomial or otherwise), one can extend the discussion here considering a single index model such as a (multinomial) Probit, a logistic or some other discrete variable model, which converts an equation such as (4) into a discrete measure.

this distribution[8] and estimate the parameters that characterize it.

Finally, we notice that the way equation (4) is written, assumes that each measure is affected exclusively by one factor. Such a system is referred to as *dedicated* system. While it is possible to relax a bit this assumption, having some measures being determined by more than one factor, identification requires a set of exclusion restrictions.

3.2 Establishing a metric

Because the latent factors are not observable and have no naturally defined metric, we need to impose normalizing restrictions to set their location and scale (Anderson and Rubin (1956)). This issue is analogous to the discussion on *anchoring* in Cunha et al. (2010). Although the issue is relatively simple when considering children of a single age, the study of the *process of child development* makes it considerably more problematic.

As discussed by Agostinelli and Wiswall (2016), normalizing the mean of the factor to be zero in each period can be a non-innocuous assumption when interested in capturing the dynamic growth of factors over time, as it would restrict the set of admissible production functions to those that are mean stationary. Moreover, valid comparison of technology parameters over time, which is key for the identification of *window of opportunities*, requires that the scale of the latent factors is consistently expressed in the same metric over ages (see Agostinelli and Wiswall (2025)). Expressing the scale of the latent factor arbitrarily, by normalizing the loading on *different* measures in each period, would imply that the technology parameters are themselves expressed in different metrics over time and therefore impossible to compare.

One straightforward solution to both issues would be available if one had at least one measure for all available time periods, and if the parameters of the measurement system do not vary for such a variable. The location and scale of the factor could then be expressed in terms of this measure, which we label as measure 1, imposing $\alpha^1_{k,t} = 0$ and $\lambda^1_{k,t} = 1 \; \forall t = 1...T$. These assumptions are necessary to identify, for each age, the mean and the variance of the latent factor $\theta_{k,t}$. Alternatively, as is sometimes done, one can leave $\alpha^j_{k,t}$ unrestricted for every t and constrain the mean of the distribution of $\theta^j_{k,t}$ to zero. Although these two assumptions are observationally equivalent, the latter makes analyzing child development over time difficult, in particular when assessing child growth.

Normalizing one of the $\alpha^j_{k,t}$ to 0 and one of the $\lambda^1_{k,t}$ to 1 for all periods allows identification of the system's parameters, including the mean and the variance of the latent factor(s). If this identification assumption is plausible, both the means and the standard deviations of

[8] Log-normality is often assumed. Alternatively researchers have used more flexible specifications, such as a mixture of log normal. See, for instance, the discussion in Cunha et al. (2010)

the estimated latent factor are now comparable over time, as they are expressed in terms of this normalizing measure. However, we need to address the availability of a measure that covers all the ages considered and, if such a measure is available, whether these restrictions are valid.

Unfortunately, while in our data we do have one measure that covers all the ages considered for health, socio-emotional skills and parental investments, we do not have it for cognitive skills. Furthermore, the assumption of a constant factor loading and intercept for all ages is a strong one. Not only does such an assumption define the scale of the estimated factors, but it also states that the *saliency* of such measure does not change with age.

To address these issues, we exploit the fact that there is at least one time period where we have some "overlap" between different measures to express the location and scale of the cognitive factor consistently over time. In section 3.3, we provide the details of this procedure with the help of an example.

3.3 Latent factors' scale and location when no measure is available for all ages

Our data includes two sets of assessments used to measure children's cognitive skills over time: the first set is only appropriate at younger ages (0 to 3), while the second set is only appropriate for older children (3+). Therefore, the simple normalization discussed above cannot be implemented. In this section, we propose an approach that addresses the unavailability of measures spanning all ages in our sample. Moreover, our procedure alleviates the concern that the factor loadings for a given measure may differ across very distant ages.

The basic intuition behind our procedure is the following. For any two adjacent ages we need a measure that is available at both ages. We then assume that the intercepts and the slopes for this common measure are the same at *two adjacent ages*. We can derive a mapping between the location and scale parameters (the α_s and λ_s in equation (4)) of the different ages and impose these restrictions in subsequent periods, so to set the location and scale of the latent factors in a common location and scale over time.

The approach can be illustrated with a simple example, where we consider 3 time periods and four measures for a generic latent factor θ: m^{a_1}, m^{a_2},m^{b_1}, m^{b_2}. To explain our scaling procedure, we only need one latent factor, although the same logic applies when multiple factors are considered.

We assume that, at age $t = 1$, m^{a_1} and m^{a_2} are available. At age $t = 2$, measures m^{a_1} and m^{b_1} are available, while at age $t = 3$, m^{b_1} and m^{b_2} are available. At this point, the distribution of the (log) latent factor is left unspecified with mean μ_θ and variance σ^2_θ. With these assumptions, the measurement system can be expressed as follows:

$$
t = 1 \quad \begin{cases} m_1^{a_1} = \alpha_1^{a_1} + \lambda_1^{a_1}\theta_1 + \epsilon_1^{a_1} \\ m_1^{a_2} = \alpha_1^{a_2} + \lambda_1^{a_2}\theta_1 + \epsilon_1^{a_2} \end{cases}
$$

$$
t = 2 \quad \begin{cases} m_2^{a_1} = \alpha_2^{a_1} + \lambda_2^{a_1}\theta_2 + \epsilon_2^{a_1} \\ m_2^{b_1} = \alpha_2^{b_1} + \lambda_2^{b_1}\theta_2 + \epsilon_2^{b_1} \end{cases} \tag{5}
$$

$$
t = 3 \quad \begin{cases} m_3^{b_1} = \alpha_3^{b_1} + \lambda_3^{b_1}\theta_3 + \epsilon_3^{b_1} \\ m_3^{b_2} = \alpha_3^{b_2} + \lambda_3^{b_2}\theta_3 + \epsilon_3^{b_2} \end{cases}
$$

In this example, no measure is available throughout the 3 periods considered. However, measures m^{a_1} and m^{b_1} provide, with an intuitive assumption, a way to construct a metric for the factors that can be used in the 3 periods. In particular, if we normalize the intercepts and factor loadings of measure m^{a_1} at age 1 and 2 to 0 and 1, that is, in the system above, we assume $\alpha_1^{a_1} = \alpha_2^{a_1} = 0$ and $\lambda_1^{a_1} = \lambda_2^{a_1} = 1$, we can use the estimates of the intercept and factor loading of $m_2^{b_1}$, that is $\alpha_2^{b_1}$ and $\lambda_2^{b_1}$, to express the location and scale of the latent factor at age 3 in the same metric. We can do so by imposing, in period $t = 3$ the normalization $\alpha_3^{b_1} = \alpha_2^{b_1}$, *and* $\lambda_3^{b_1} = \lambda_2^{b_1}$, rather than the standard normalization which imposes an intercept to 0 and a slope to 1. This normalization imposes that the location and scale coefficients pertaining to this measure are constant across time periods. We can then identify the parameters of the factors' joint distribution (μ_θ and σ_θ^2). Our procedure guarantees that these are expressed in the same metric and are, therefore, comparable over time.

Moreover, this approach naturally accommodates the possibility that the informativeness of a given measure may change—sometimes substantially—as children age. A measure that provides useful information at 12 months may be largely uninformative by 48 months; the ability to say "mama," for example, illustrates this point. This may help explain why, in practice, certain measures are administered at specific ages. The overlapping procedure we propose handles well this practice. The key requirement is not that the parameters of a measure remain constant across all ages, but rather that they remain constant across adjacent periods during which the measure is actually administered.

In our example, we assume the loading factor for measure m^{a_1} remains constant between ages 1 and2, that is $\lambda_1^{a_1} = \lambda_2^{a_1}$ — which is equivalent to say that measure m^{a_1} is equally

informative about the factor θ at ages 1 and 2. Analogously, we assume that $\lambda_2^{b_1} = \lambda_3^{b_1}$ — that is measure m^{b_1} is equally informative at ages 2 and 3. We make no assumption about $\lambda_3^{a_1}$ or $\lambda_1^{b_1}$, which need not equal the values at adjacent periods, and which may, in practice, differ substantially. An analogous reasoning applies to the intercepts α.

More formally, imposing $\alpha_2^{b_1} = \alpha_3^{b_1}$ and $\lambda_2^{b_1} = \lambda_3^{b_1}$ yields:

$$E(\theta_3) = E(\theta_2) + \frac{E(m_3^{b_1}) - E(m_2^{b_1})}{\lambda_2^{b_1}} \tag{6}$$

The factor mean at age 3 equals that at age 2 plus the change in measure m^{b_1} between the two periods, scaled by its informativeness. This example generalizes to additional periods and measures: wherever two measures overlap across adjacent ages, they can be used to carry the metric forward. The crucial identifying assumption is that the factor loadings and intercepts of the bridging measures are stable across the adjacent periods in which they are used.

This example can be generalized to additional time periods and additional measures. What we want to make clear is that the availability of two measures in a given time period allows us to "convert" one to the other. We can then impose these conversions in different periods and express the scale and location of the latent factors in a way that is consistent throughout the periods considered. The resulting factors can then be compared over time to establish the rate of growth over time and to derive estimates of the parameters governing the growth process. The crucial identification assumptions, again, is that the factor loadings and intercepts of some of the measures available at adjacent ages are constant. In what follows, we will be using the TVIP to express the scale and the location of the cognitive factor.

3.4 Latent factors distributions and measurement system parameters

Having established a metric for the latent factors under consideration, the next step is to estimate their distributions along with the parameters of the measurement system (intercepts and factor loadings). While these could in principle be estimated non-parametrically, we instead adopt flexible parametric assumptions. Specifically, to avoid placing rigid constraints on the production technology for human capital, the distribution of inputs must be estimated with sufficient flexibility. Common assumptions like joint normality, while convenient, impose strong restrictions in practice — forcing a linear relationship among conditional means and ruling out any complementarity between inputs. A flexible production technology therefore requires equally flexible assumptions about the joint distribution of those inputs.

We follow Attanasio et al. (2020a)'s approach, allowing for flexible distributions of

the latent factors and measurement errors and estimating the parameters of these distributions, together with those of the measurement system, from the observed data. Specifically, we assume that the latent factors follow a mixture of two log-normal distributions, while the measurement errors are normally distributed with mean zero and a diagonal variance-covariance matrix.

These assumptions, combined with the measurement model, imply that the observed measures follow a mixture of normal distributions. We therefore estimate the parameters of the joint distribution of measurements by maximum likelihood using the EM algorithm. From these estimates, we then recover the parameters of the joint distribution of the latent factors, the measurement-error variances, the factor loadings, and the intercepts by minimum distance. Following Attanasio et al. (2020a), we augment the measurement system with additional equations representing the control variables and the instruments used in the investment equation, the latter used to address the endogeneity of parental investment, which we discuss in the following.

4 AEIOTU: Available data and latent factors.

The data we are using to estimate the properties of the process of child development were collected for an RCT designed to evaluate the impact of an innovative intervention aimed at providing high quality child care to young children in the city of Santa Marta in Colombia.[9] This data set is unique in some important dimensions. In particular, rich data on child development and its drivers were collected for several years at an annual frequency. Although the data do not form a balanced panel, we observe children from 6 months to 7 years of age at annual frequency. We are not aware of other data sets that allow the identification and measurement of several dimensions of child development at such high frequency.

In this section, we start by providing some information on the intervention evaluated by the RCT that collected the data we are using, and on the population it targeted. We then discuss the measurement system we use to estimate the factors of interest and the items that are used to identify such factors.

4.1 The intervention, its evaluation, and the population studied

The intervention we are considering was designed to offer high-quality center-based early child care and education to socioeconomically disadvantaged children under age five in two under-served communities in northern Colombia.

[9]The impact of the intervention is the subject of different studies (see Nores et al. (2019), Bernal et al. (2024)).

The program, named *aeioTU*, was run by a Colombian NGO (*Fundación Carulla*) and delivered through a public–private partnership linked to Colombia's *De Cero a Siempre* early childhood strategy. It was centered in centers providing integrated child care and education, including nutrition and health components. The program offered full-day care (about 9 hours per day) for 11 months per year, in classrooms characterized by low child-to-teacher ratios (4:1 for infants and 6:1 for toddlers), relatively high teacher qualifications, extensive pre- and in-service training, and pedagogical coordination. It engaged families through workshops and regular communication, and relied on data systems to monitor and improve quality.

The education curriculum was inspired by the celebrated *Reggio Emilia* approach (see Malaguzzi (1993)), emphasizing project- and play-based learning, rich adult–child interactions, learning across developmental domains, and a balanced mix of teacher-directed and child-initiated activities. The program also included health and nutrition services ensuring that children received approximately 70% of their daily nutritional requirements through meals at the center, along with regular monitoring by on-site nutritionists and nutritional supplementation during holidays.

The evaluation was designed as a randomized controlled trial that exploited the excess demand for places in two newly built *aeioTU* centers scheduled to open in 2010. Eligible children were identified through a door-to-door census conducted jointly with *aeioTU* and community leaders. Children were individually randomized via an oversubscription lottery to receive an offer of admission to an *aeioTU* center or to a control group that did not receive an initial offer. As a consequence, differences in outcomes between treated and control children measure the impact of an offer of *aeioTU* relative to the mix of care arrangements otherwise available in the community—including home care, other center-based care, later transition to primary school, and, for some children, later crossover into *aeioTU* as places became available.

A sample of children was observed over several years, a random subset of whom was admitted to *aeioTU*. In Table 1, we report some summary statistics for the sample we use, computed from the baseline data. At baseline, the average age of the children involved was 25 months, with a majority of black children. About 28% of the mothers were single and 23% were working. The average number of family members was over 5, and in slightly less than 70% of households the father of the child was living at home. Less than 40% of mothers or fathers had completed secondary education or above. Finally, about 34% of the households considered were beneficiaries of the Colombian CCT program *Familias en Acción*.

Table 1: Sample Summary Statistics

	Mean	SD
Age in months	25.03	12.39
Race: black (%)	61.60	48.66
Mother is single (%)	27.64	44.74
Mother works (%)	23.77	42.59
Mother secondary complete or above (%)	36.32	48.12
Childcare use at baseline (%)	17.36	37.89
Father lives at home (%)	69.06	46.25
Father secondary complete or above (%)	39.94	49.00
Children's books at home	1.43	2.58
Household size	5.35	2.01
CCT beneficiary (%)	34.34	47.51
Number of children	1060	

Notes: This table reports the summary statistics for the analysis sample at baseline.

4.2 Available measures and the measurement system

The data set includes several measures available for each of the latent factors that enter the production functions we will be estimating. We use the factor models described in Section 3 to summarize all the information available. In particular, an exploratory factor analysis on the measures and *a-priori* considerations lead us to consider three latent factors of development: (i) cognition and language, (ii) socio-emotional skills, and (iii) health.

Of course, a finer decomposition of child development and, consequently, a larger number of factors could be possible. For instance, one could imagine considering different dimensions of cognitive development, including language, emergent math, executive functions and other types of cognitive skills. Or one could decompose what we call socio-emotional skills into 'internalizing skills' (focus, motivation, drive and grit) and 'externalizing skills' (sociability, the ability to work and communicate with others).[10] However, given the data available and the exploratory analysis we performed, the decomposition we adopt seems reasonable and feasible. These three factors have been used extensively in the literature on child development, although, to the best of our knowledge, they have not previously been modeled jointly.

[10]The areas measured by the Ages and Stages Questionnaire (ASQ SE) are self-regulation, compliance, social-communication, adaptive functioning, autonomy, affect, and interaction.

The available measures lead us to specify a *dedicated* measurement system across three dimensions of child development — cognition and language, socio-emotional skills, and health — as well as parental investment. Table 2 lists the observable variables used as markers for each latent factor, along with, for each marker, the share of its variance attributable to the underlying factor and the share due to measurement error.

The other issue already mentioned is that the measures used to assess children change with age, and no single cognitive measure is used throughout the ages over which we observe children in the sample. However, the evaluation was designed with overlap across some ages, so that for each adjacent pair of ages we can follow the normalization and anchoring strategy discussed in Section 3.3.

The cognitive measures change with the age of the children, as is normal practice. At ages 1 to 3, the sample children were administered various subscales of the Bayley Scales of Infant Development, 3rd edition (Bayley, 2006, 2005), a test that is considered the gold standard for infant and toddler ages. The quality of the Bayley scales is reflected in the high signal-to-noise ratios corresponding to these measures that we report in Table 2. We notice, in particular, the high signal-to-noise ratio for the cognitive scale and the two language scales (expressive and receptive). Starting at age 3, children are also administered the Test de Vocabulario en Imágenes (TVIP, Padilla et al. (1986)). From age 4, children are also administered the Woodcock-Muñoz, 3rd edition (Muñoz-Sandoval et al. (2005)) tests (specifically, the applied problems and letter-word subtests) and the Early Literacy Skills Assessment (ELSA, DeBruin-Parecki (2005)).

To apply the methodology discussed in Section 3, the TVIP plays an important role, as it is observed at age 3 together with the Bayley (used for children up to that age) and at age 4 together with the Woodcock-Muñoz, which is used after age 3. It therefore defines our metric.

For socio-emotional skills, children are administered five subscales of the Ages and Stages Questionnaire: Socio-Emotional, 1st edition (ASQ:SE, Squires et al. (2002)) at ages 1 and 2, and six subscales at ages 3 to 6. Unfortunately, the signal-to-noise ratio for several of these subscales is relatively low. This relatively low information content is one of the reasons that prevents us from decomposing socio-emotional skills into separate internalizing and externalizing factors.

For health, the three measures included are weight, height, and arm circumference. It is interesting to note that in the first few years of life, weight and height are very informative, while arm circumference starts to be somewhat informative at age 4 and becomes more important past age 5.

Finally, for parental investments, we have several measures from the FCI test. As also

Table 2: Share of total variance due to signal and noise

	% signal	% noise		% signal	% noise
Child cognitive skills			*Child socio-emotional skills*		
Bayley cognition age 1	0.91	0.09	ASQ selfregulation age 1	0.12	0.88
Bayley expressive age 1	0.58	0.42	ASQ communication age 1	0.01	0.99
Bayley receptive age 1	0.61	0.39	ASQ adaptive functioning age 1	0.22	0.78
Bayley fine motor age 1	0.86	0.14	ASQ compliance age 1	0.63	0.37
Bayley gross motor age 1	0.9	0.1	ASQ interaction age 1	0.01	0.99
Bayley cognition age 2	1†	0†	ASQ selfregulation age 2	0.11	0.89
Bayley expressive age 2	0.54	0.46	ASQ communication age 2	0.15	0.85
Bayley receptive age 2	0.44	0.56	ASQ adaptive functioning age 2	0.02	0.98
Bayley fine motor age 2	0.6	0.4	ASQ compliance age 2	0.24	0.76
Bayley gross motor age 2	0.6	0.4	ASQ interactiona age 2	0.09	0.91
TVIP age 3	1†	0†	ASQ selfregulation age 3	0.18	0.82
Bayley cognition age 3	0.89	0.11	ASQ communication age 3	0.2	0.8
Bayley expressive age 3	0.6	0.4	ASQ adaptive functioning age 3	0.19	0.81
Bayley receptive age 3	0.38	0.62	ASQ affection age 3	0.16	0.84
Bayley fine motor age 3	0.5	0.5	ASQ compliance age 3	0.34	0.66
Bayley gross motor age 3	0.43	0.57	ASQ interaction age 3	0.11	0.89
TVIP age 4	0.71	0.29	ASQ selfregulation age 4	0.2	0.8
Woodcock-Munoz age 4	0.26	0.74	ASQ communication age 4	0.26	0.74
ELSA age 4	0.38	0.62	ASQ adaptive functioning age 4	0.15	0.85
TVIP age 5	0.27	0.73	ASQ affection age 4	0.5	0.5
Woodcock-Munoz age 5	0.32	0.68	ASQ compliance age 4	0.25	0.75
ELSA age 5	0.36	0.64	ASQ interaction age 4	0.3	0.7
TVIP age 6	0.43	0.57	ASQ selfregulation age 5	0.15	0.85
Woodcock-Munoz age 6	0.54	0.46	ASQ communication age 5	0.64	0.36
ELSA age 6	0.23	0.77	ASQ adaptive functioning age 5	0.14	0.86
TVIP age 7	0.42	0.58	ASQ affection age 5	0.42	0.58
Woodcock-Munoz age 7	0.44	0.56	ASQ compliance age 5	0.23	0.77
ELSA age 7	0.15	0.85	ASQ interaction age 5	0.34	0.66

continues...

Notes: †0 and 1 reflect rounding. ASQ is the Ages and Stages Questionnaire, TVIP is the "Test de Vocabulario en Imágenes Peabody" — the Spanish-language adaptation of the Peabody Picture Vocabulary Test (PPVT), ELSA is the Early Literacy Skills Assessment (DeBruin-Parecki, 2005).

Share of total variance due to signal and noise (continued)

	% signal	% noise		% signal	% noise
Child socio-emotional skills			*Parental investments*		
ASQ selfregulation age 6	0.4	0.6	Child books ages 1-2	0.05	0.95
ASQ communication age 6	0.28	0.72	Number of play materials ages 1-2	0.41	0.59
ASQ adaptive functioning age 6	0.17	0.83	Books at home ages 1-2	0.93	0.07
ASQ affection age 6	0.33	0.67	Toys ages 1-2	0.08	0.92
ASQ compliance age 6	0.4	0.6	At least 3 child books ages 1-2	0.08	0.92
ASQ interaction age 6	0.29	0.71	Learning materials (score) ages 1-2	0.07	0.93
Child health			Variety (score) ages 1-2	$0^{\dagger}$	$1^{\dagger}$
Weight age 1	0.67	0.33	Child books ages 3-4	0.15	0.85
Arm circumference age 1	0.24	0.76	Books at home ages 3-4	0.06	0.94
Height age 1	0.79	0.21	Number of play materials ages 3-4	0.84	0.16
Weight age 2	0.71	0.29	Toys to learn colors ages 3-4	0.73	0.27
Arm circumference age 2	0.09	0.91	Games ages 3-4	0.52	0.48
Height age 2	0.69	0.31	Soft toys ages 3-4	0.69	0.31
Weight age 3	0.88	0.12	Toys to learn numbers ages 3-4	0.59	0.41
Arm circumference age 3	0.1	0.9	At least 10 child books ages 3-4	0.73	0.27
Height age 3	0.91	0.09	Mother activities with child ages 3-4	0.08	0.92
Weight age 4	0.73	0.27	Child books ages 5-6	0.13	0.87
Arm circumference age 4	0.25	0.75	Books at home ages 5-6	0.05	0.95
Height age 4	0.9	0.1	Number of play materials ages 5-6	0.82	0.18
Weight age 5	0.67	0.33	Toys to learn shapes ages 5-6	0.57	0.43
Arm circumference age 5	0.24	0.76	Puzzles ages 5-6	0.48	0.52
Height age 5	0.98	0.02	Toys for music ages 5-6	0.07	0.93
Weight age 6	0.81	0.19	Games ages 5-6	0.6	0.4
Arm circumference age 6	0.51	0.49	Soft toys ages 5-6	0.62	0.38
Height age 6	0.67	0.33	Toys to learn numbers ages 5-6	0.66	0.34
Weight age 7	0.79	0.21	At least 10 child books ages 5-6	0.28	0.72
Arm circumference age 7	0.55	0.45	Toys to learn names ages 5-6	0.46	0.54
Height age 7	0.64	0.36	Musical instruments ages 5-6	0.48	0.52

Notes: †0 and 1 reflect rounding

shown in Table 2, these measures are good markers of parental investment and contain a substantial amount of information. All the items we use refer to material rather than time investment, with the exception of "Mother activities with child ages 3-4."

5 The evolution of child development: descriptive evidence

Estimating the measurement system across all latent factors and ages recovers the distribution of each factor and how it evolves over time. For identification, we follow the procedure described in Section 3.3 for each factor, restricting the slope and the intercept corresponding to one of the available measures so to maintain a coherent metric, while allowing the factors' means and standard deviations to vary with age. Under the assumptions we have made, this protocol allows us capturing growth over time in the means of the latent factors considered, as well as changes in their distribution.

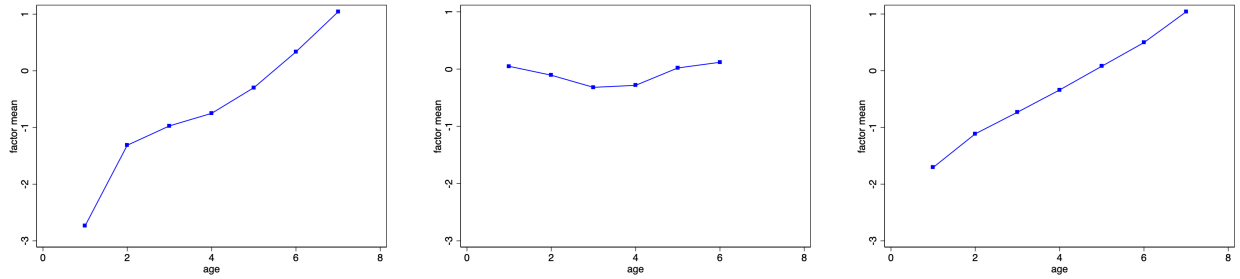
Table 3: Evolution of the latent factors mean and standard deviations with age

	Mean						
	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
Cognitive factor	-2.73	-1.52	-0.10	-0.72	-0.25	0.33	1.04
Socio-emotional factor	0.048	-0.107	-0.319	-0.284	0.02	0.118	-
Health factor	-1.71	-1.113	-0.732	-0.339	0.079	0.496	1.039

	St Dev.						
	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
Cognitive factor	0.55	0.36	0.26	0.29	0.41	0.60	0.60
Socio-emotional factor	0.45	0.35	0.43	0.51	0.47	0.65	-
Health factor	0.39	0.29	0.32	0.34	0.39	0.58	0.73

In Table 3, we report how the means and standard deviations of the three developmental factors vary with age.[11] Figure 1 plots the factor means against age. As expected, the mean of the cognitive factor grows with age. The socio-emotional factor, by contrast, is relatively flat: it dips after age 1 and recovers after age 4, but the changes are not substantial. Finally, the estimated means for health grow monotonically with age.

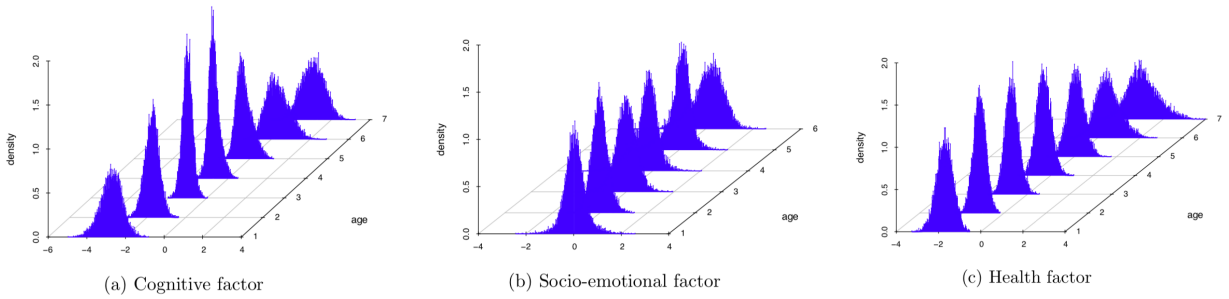
[11] For socio-emotional skills, estimation is limited to age 6, since no valid measures exist at age 7.



(a) Cognitive factor (b) Socio-emotional factor (c) Health factor

Figure 1: Latent Factors' Means Over Time

The means of the three child development factors are not the only moments of interest. Having estimated the full distribution for our sample and its evolution over age, we can examine additional features of the data. First, in the bottom panel of Table 2, we report how the standard deviations of the three factors. For cognition, the pattern of standard deviations is U-shaped, being higher at the beginning and end of the interval considered. The standard deviation for socioemotional skills, instead, is relatively flat, with the exception of a marked increase at age 6. Finally, for the health factor, the standard deviation is also relatively flat, except for the last two ages when it increases.



(a) Cognitive factor

(b) Socio-emotional factor

(c) Health factor

(a) Cognitive factor (b) Socio-emotional factor (c) Health factor

Figure 2: Latent Factors' Distributions Over Time

Figure 2 presents the age profiles of the estimated distributions for the three factors. For these and for all ages considered, the estimated distributions are unimodal. It is worth noting that this is not imposed by the estimation, as we assume that the distribution of the factors is given by a mixture of normal distributions. In addition, for the three factors, the dispersion of the distribution increases considerably at age 6. Interestingly, for the cognitive factor, the distribution at age 3 and 4 is considerably tighter.

6 Parental investment and the dynamics of child development

Having estimated the joint distribution of all the variables included in the model through the measurement system described above, we draw vectors of random variables from this distribution. Because these are based on the estimated joint distribution, the resulting synthetic data reproduce the correlation observed in the original data. We then use this synthetic data to estimate the parameters of the structural model using a control function approach discussed below.

An important challenge in estimating the production function is that some of its arguments are endogenous. In particular, as mentioned earlier, parents may adjust their investments in response to their child development over time. In addition, the latent factors may be correlated with other unobserved factors also relevant to child development. These complicate the identification of the causal link between parental investments and child development.

In what follows, we use the approach in Attanasio et al. (2020a). In particular, we first estimate an *investment function*, which represents the reduced form of a model that relates parental investment to (some of) its determinants.

The residual terms of this equation are used to compute a *control function* that addresses the endogeneity of investment in the production function of human capital. The parameters of the production function characterize the role of parental investment and other inputs in the development process, the dynamics of skill accumulation, and the interactions across domains at different ages. The identification assumption is that the investment function includes some variables that are important drivers of investment but do not directly enter the production function.

The investment equation, its residuals, and the production function for each of the three latent factors – augmented with the corresponding control functions – are estimated jointly on the synthetic data described above. To obtain standard errors, we bootstrap this entire procedure: the variability of the resulting coefficients across bootstrap replications is used to construct the standard errors and confidence intervals reported in the tables.

6.1 Investment functions

We first use our measurement system estimates to recover the investment functions (3) introduced in Section 2.2.3. At each age, parental investment is modeled as a function of the child's developmental status (captured by the three latent factors), parental education the number and gender of children in the household, household income, household size, and an indicator for the presence of the father. The investment measure used here only captures

material inputs in the home and therefore excludes time investment and any investment children receive at child care or elementary school. Results are reported in Table 4.

Turning first to how parents react to their child's developmental status, the coefficient on cognition in the investment equation is not significantly different from zero at ages 2, 3, and 5, and is marginally significant and negative at age 4 — suggesting that children with lower cognitive development receive slightly higher parental investment at that age. At ages 6 and 7, however, the coefficient is positive and strongly significant: at these later ages, parents invest more in children exhibiting higher cognitive skills. The coefficients on current socioemotional skills are never statistically different from zero, with an exception at age 6, where the effect is positive but only significant at the 10% level. The coefficient on health is positive and marginally significant at age 5, but otherwise not different from zero at all other ages.[12]

Parental education has a strong, positive effect on investment that is statistically significant at all ages, with the coefficient increasing with age and becoming particularly large at ages 6 and 7. The number of children in the household is negatively associated with parental investment at age 5. We find no statistically significant effect of child gender on parental investment. The presence of the father has a strong positive effect on investment, while total household size has a negative effect.

The income coefficient is positive and statistically significant throughout. This relationship is particularly important for identification, as income provides the source of exogenous variation used to identify the effect of parental investment in the production function. While the exclusion restriction that income affects child development only through parental investment may be questioned, at least, income is not a weak instrument.

6.2 The production functions of human development

Having estimated the investment function, we now turn to the production function for the three developmental factors included. We explore several alternatives, including Translog and CES specifications, but in most cases cannot reject the restrictions implied by the Cobb-Douglas; where deviations were found, they were only marginally significant and had limited quantitative implications. We therefore conclude that the Cobb-Douglas provides a good approximation to the data. For each of the three domains of child development — cognition, socioemotional skills, and health — we present a specification that includes one lag of all three relevant domains; that is, we allow cognition, socioemotional skills, and health

[12]We also tried including a treatment indicator. However, consistent with the evidence in Bernal et al. (2024), *aeioTU* did not appear to have any impact, positive or negative, on parental investment. We therefore exclude it from the investment equation; its inclusion does not affect the results we obtain.

Table 4: Investment equations

	Investment equations					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
Intercept	0.11	0.071	-0.264	-0.103	-0.155	-0.266
	0.206	0.165	0.097	0.103	0.085	0.077
	[-0.254,0.417]	[-0.221,0.327]	[-0.394,-0.071]	[-0.201,0.115]	[-0.252,0.028]	[-0.352,-0.092]
Cognition t	0.084	0.026	-0.099	0.115	0.147	0.093
	0.078	0.087	0.054	0.091	0.077	0.062
	[-0.045,0.205]	[-0.081,0.202]	[-0.198,-0.01]	[-0.014,0.268]	[0.035,0.279]	[-0.01,0.179]
Socio-emotional t	-0.231	-0.05	-0.027	-0.01	0.11	0.022
	0.29	0.114	0.036	0.051	0.07	0.049
	[-0.392,0.491]	[-0.163,0.199]	[-0.088,0.023]	[-0.071,0.088]	[0.027,0.25]	[-0.039,0.123]
Health t	0.045	0.201	0.085	0.118	0.018	0.043
	0.125	0.103	0.067	0.066	0.054	0.04
	[-0.155,0.248]	[-0.024,0.302]	[0,0.217]	[-0.001,0.216]	[-0.072,0.114]	[-0.016,0.118]
Parental Cognition	0.062	0.059	0.233	0.229	0.35	0.357
	0.125	0.114	0.067	0.064	0.092	0.095
	[0.011,0.405]	[0.007,0.377]	[0.144,0.345]	[0.14,0.341]	[0.191,0.489]	[0.192,0.493]
Number of Children	-0.259	-0.296	-0.314	-0.317	-0.329	-0.347
	0.115	0.083	0.058	0.059	0.072	0.071
	[-0.445,-0.057]	[-0.404,-0.143]	[-0.391,-0.197]	[-0.375,-0.177]	[-0.395,-0.157]	[-0.409,-0.176]
Gender	0	0.003	0.043	0.035	0.033	0.029
	0.034	0.031	0.026	0.026	0.031	0.032
	[-0.055,0.053]	[-0.043,0.055]	[-0.01,0.077]	[-0.015,0.069]	[-0.029,0.081]	[-0.04,0.072]
Income	0.185	0.197	0.233	0.232	0.175	0.179
	0.067	0.061	0.044	0.043	0.043	0.044
	[0.064,0.274]	[0.075,0.268]	[0.15,0.29]	[0.14,0.277]	[0.082,0.227]	[0.085,0.232]
Household size	-0.293	-0.295	-0.106	-0.091	-0.056	-0.063
	0.077	0.065	0.044	0.043	0.067	0.067
	[-0.414,-0.168]	[-0.39,-0.184]	[-0.166,-0.015]	[-0.146,-0.002]	[-0.156,0.057]	[-0.161,0.052]
Father present	0.149	0.145	0.126	0.135	0.104	0.12
	0.04	0.04	0.023	0.024	0.03	0.028
	[0.082,0.211]	[0.087,0.208]	[0.076,0.15]	[0.065,0.139]	[0.03,0.124]	[0.047,0.133]

Notes: This Table shows the estimates of the investment functions at different ages. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets.

at age $t+1$ to depend, among other variables, on cognition, socioemotional skills, and health at age t.

We also considered specifications with additional lags allowing for cross-domain effects — for instance, the influence of current socioemotional skills and health on future cognitive development — which would relax the Markov assumption commonly considered in the literature. We found few significant effects of this type, however, and to preserve precision and avoid cluttering the tables, we report these results in the Appendix. Each table reports production function estimates for ages 2 through 7, except for socioemotional skills, which, as mentioned above, is estimated only up to age 6 due to data limitations. Each table has two panels: the left ignores investment endogeneity, while the right applies the control function approach described above, allowing us to directly assess how much accounting for endogenous investment matters in practice.

Our choice of instruments for the control function approach is constrained by the available data. The only feasible option is to take the model outlined in Section 2.2.3 at face value and instrument parental investment with parental income, which enters the budget constraint but not the production function directly. As mentioned, this strategy has limitations, since income may correlate with inputs omitted from the production function. That said, using a different Colombian dataset, Attanasio et al. (2020b) instrument parental investment with local prices and conflict exposure — both of which vary across localities — and find their results largely robust to using household resources alone as an instrument, lending some support to our approach. These results show how the parameters of the production function change as children age, and can be used to pinpoint when and where interventions might be most effective.

6.2.1 The process of cognitive development at different ages

We start with a discussion of the results for the production function of cognitive development, reported in Table 5. As mentioned above, the left panel of the table contains the OLS estimates, which ignore the endogeneity of parental investment, while the right panel reports estimates of the production function that include the control function derived from the investment function in Table 4.

Both panels report estimates at each age between 2 and 7. The estimation of a production function for child development at this frequency is an important novelty of this paper, as it helps identify the ages and periods in which investment might be particularly impactful, as we discuss in the next section. It also identifies the ages at which the current state of development becomes particularly relevant for subsequent development.

The coefficients on parental investment are significantly different from zero and quanti-

tatively important only between ages 2 and 4. At ages 3 and 4, treating parental investment as endogenous roughly doubles the estimated coefficients — a finding consistent with Attanasio et al. (2020b), Attanasio et al. (2020a), and Attanasio et al. (2017). The coefficient on parental investment at age 5 and older is statistically indistinguishable from zero, regardless of whether parental investment is treated as endogenous and regardless of the dynamic structure imposed on the production function. Importantly, this near-zero estimate does not reflect a loss of precision but a reduction in the point estimates.

Parental education does not appear to play a significant role in the cognitive production function, at least once parental investment is controlled for. It does, however, strongly predict parental investment itself, suggesting that its influence on child development operates primarily through the investment channel — a finding in line with Attanasio et al. (2020b). Finally, the intervention, which altered the quality of care children received in their nurseries, had a positive and statistically significant, if moderate, effect at ages 4 and 5. The point estimate at age 7 is of comparable magnitude but falls short of statistical significance due to reduced precision.

Turning to the dynamics of the process, the persistence of cognitive development changes considerably with age, as shown by the coefficient on lagged cognition. Up to and including age 4, this coefficient is significantly different from zero, but the process is far from highly persistent, with coefficients around 0.3 — a result that holds in both the OLS and CF estimates. At age 5, however, the coefficient on lagged cognition increases considerably and is close to 1.

The coefficients on lagged socioemotional skills are never significant in the regressions for cognitive development. The coefficient on lagged health status, by contrast, is significant and quantitatively important at some ages: lagged health matters for cognitive development particularly at ages 2 and 4, and is also marginally significant — though considerably smaller — at age 3. As we will see in the next section, these results turn out to be important for the dynamic targeting of investment and interventions.

6.2.2 The process of socioemotional skills development at different ages

We now turn to the results for the socioemotional skills production function, reported in Table 6. The table follows the same structure as the one used for cognitive skills, with one notable difference: the analysis of socioemotional skills is restricted to ages 1 through 6, reflecting the availability of suitable measures for this domain.

Before discussing the estimates, it is important to again acknowledge that the quality of the available measures for socioemotional skills is considerably lower than for the other developmental domains included, as shown in Table 2. and discussed above. This relatively

poor measurement quality implies that estimates for this domain are likely to be less precise and should be interpreted with caution.

Starting with the dynamics of the process, neither lagged cognitive skills nor lagged health are ever statistically significant. Regarding lagged socioemotional skills, persistence does not increase monotonically with age, in contrast to the pattern observed for cognitive skills. The coefficient on lagged socioemotional skills is approximately 0.3 across most ages, with the exception of age 3, where it is estimated at 0.7, albeit with considerable imprecision.

Turning to the investment coefficients, the control function is statistically significant in almost all specifications, providing evidence of investment endogeneity. Its negative sign indicates that treating investment as endogenous substantially increases the estimated magnitude of its effect: at age 4, for example, the point estimate of the investment coefficient rises from 0.135 to 0.506 once endogeneity is accounted for.

Notably, unlike cognitive skills — where the investment coefficient became effectively zero at age 5 and beyond — the effect of parental investment on socioemotional skills does not decline appreciably with age. At age 5, the coefficient remains sizeable at 0.654, and although it falls to 0.354 at age 6, it remains statistically significant.

Regarding the intervention, a significant impact is observed at ages 4 and 5. Finally, when parental investment is treated as endogenous, parental education is not statistically significant in the production function — a result that mirrors the finding for cognitive skills.

6.2.3 The health process at different ages

Finally, we discuss the estimates of the production function for the health domain of child development, reported in Table 7. As with the other two domains, the left panel treats parental investment as exogenous, while the right panel uses the CF approach to control for its possible endogeneity. As with cognition, we estimate the function for ages 2 through 7.

Starting with the dynamic properties of the process, both socioemotional and cognitive skills appear marginally significant at some ages, though the size of these coefficients is small enough that their effect is at best marginal. Lagged health status, by contrast, is very important, signaling strong persistence: the coefficient already exceeds 0.7 at age 2, and is close to 1 at older ages.

Turning to the effect of parental investment, age 2 is the only age at which it is productive — and also the only age at which it appears to be endogenous. As with our other production function estimates, the control function takes a negative sign at age 2, so its inclusion increases the coefficient on parental investment from 0.27 to 0.37. Past age 2, parental investment does not attract coefficients significantly different from zero.

Finally, neither the aeioTU intervention nor parental education have a direct effect

on the health production function, except at ages 6 and 7, when the intervention effect is marginally significant. Regarding parental education, we recall that it affects significantly parental investment, which is itself significant at age 2. Given the high persistence of the health process, impacts at this early age are likely extend into subsequent outcomes.

Table 5: Production functions for cognitive skills

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.155	-0.135	0.44	0.552	0.74	-0.29	0.027	-0.176	0.481	0.548	0.739
	0.16	0.11	0.03	0.054	0.043	0.05	0.154	0.114	0.044	0.07	0.05	0.05
	[-0.443,0.079]	[0.048,0.402]	[-0.191,-0.098]	[0.341,0.517]	[0.486,0.624]	[0.651,0.811]	[-0.52,-0.037]	[-0.097,0.299]	[-0.278,-0.145]	[0.37,0.588]	[0.456,0.626]	[0.649,0.813]
Control function	-	-	-	-	-	-	-0.22	-0.268	-0.132	0.133	-0.028	-0.061
							0.145	0.102	0.098	0.113	0.161	0.12
							[-0.423,0.036]	[-0.43,-0.127]	[-0.366,-0.053]	[-0.017,0.333]	[-0.272,0.193]	[-0.223,0.158]
AeioTu	-0.13	0.034	0.03	0.068	0.025	0.05	-0.11	0.046	0.03	0.068	0.026	0.056
	0.081	0.056	0.013	0.018	0.045	0.042	0.078	0.051	0.013	0.018	0.046	0.045
	[-0.217,0.038]	[-0.048,0.125]	[0.007,0.047]	[0.026,0.09]	[-0.048,0.099]	[-0.004,0.122]	[-0.208,0.046]	[-0.036,0.129]	[0.007,0.046]	[0.026,0.088]	[-0.051,0.093]	[-0.005,0.134]
Cognitive t	0.3	0.501	0.436	1.021	1.138	0.908	0.3	0.494	0.423	1.051	1.133	0.898
	0.067	0.081	0.039	0.085	0.059	0.048	0.067	0.082	0.042	0.1	0.072	0.054
	[0.229,0.442]	[0.449,0.692]	[0.372,0.498]	[0.879,1.141]	[1.01,1.205]	[0.805,0.954]	[0.221,0.442]	[0.421,0.678]	[0.349,0.487]	[0.901,1.207]	[0.982,1.22]	[0.79,0.969]
Socio-emotional t	0.08	0.158	0.058	-0.06	0.007	0.058	0.08	0.121	0.046	-0.043	0.002	0.053
	0.103	0.075	0.023	0.031	0.036	0.043	0.101	0.074	0.028	0.035	0.049	0.044
	[-0.007,0.335]	[0.018,0.263]	[0.008,0.085]	[-0.089,0.014]	[-0.074,0.044]	[-0.001,0.133]	[-0.014,0.326]	[-0.022,0.215]	[-0.017,0.071]	[-0.071,0.038]	[-0.1,0.063]	[-0.012,0.139]
Health t	0.39	0.238	0.293	0.006	-0.065	-0.062	0.35	0.161	0.267	0.024	-0.066	-0.066
	0.12	0.103	0.048	0.063	0.044	0.034	0.116	0.097	0.055	0.063	0.045	0.035
	[0.134,0.512]	[0.051,0.366]	[0.193,0.362]	[-0.087,0.108]	[-0.119,0.021]	[-0.095,0.011]	[0.108,0.478]	[-0.003,0.319]	[0.156,0.327]	[-0.067,0.137]	[-0.123,0.02]	[-0.103,0.009]
Parental cognition	0.02	0.013	0.102	-0.011	-0.031	0.101	0	-0.008	0.061	0.025	-0.04	0.079
	0.089	0.054	0.031	0.036	0.045	0.045	0.089	0.061	0.044	0.046	0.068	0.064
	[-0.017,0.233]	[-0.011,0.142]	[0.058,0.159]	[-0.051,0.065]	[-0.085,0.064]	[-0.008,0.142]	[-0.047,0.223]	[-0.103,0.063]	[-0.027,0.118]	[-0.023,0.121]	[-0.138,0.086]	[-0.045,0.16]
Investments	0.2	0.091	0.111	0.043	-0.05	-0.005	0.27	0.231	0.203	-0.057	-0.03	0.036
	0.085	0.068	0.041	0.042	0.053	0.045	0.12	0.103	0.092	0.103	0.145	0.097
	[-0.017,0.28]	[-0.035,0.184]	[0.061,0.191]	[-0.06,0.081]	[-0.129,0.045]	[-0.032,0.106]	[-0.032,0.376]	[0.082,0.387]	[0.134,0.418]	[-0.273,0.04]	[-0.234,0.232]	[-0.113,0.19]

Notes: This Table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).

Table 6: Production functions for socioemotional skills

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.104	0.217	0.312	0.286	0.11	-0.12	0.054	0.09	0.236
	0.303	0.408	0.07	0.064	0.079	0.326	0.409	0.104	0.108	0.089
	[-0.143,0.812]	[-0.596,0.706]	[0.1,0.327]	[0.222,0.451]	[0.158,0.407]	[-0.483,0.597]	[-0.893,0.355]	[-0.141,0.175]	[-0.102,0.251]	[0.102,0.377]
Control function	-	-	-	-	-	-0.62	-0.471	-0.529	-0.728	-0.356
						0.279	0.263	0.237	0.273	0.348
						[-0.979,-0.097]	[-1.045,-0.244]	[-0.955,-0.241]	[-1.288,-0.368]	[-0.947,0.169]
AeioTu	-0.12	0.13	-0.007	-0.115	-0.001	-0.08	0.151	-0.006	-0.117	0.017
	0.181	0.162	0.038	0.031	0.1	0.173	0.156	0.037	0.031	0.106
	[-0.375,0.234]	[-0.305,0.219]	[-0.072,0.05]	[-0.175,-0.072]	[-0.195,0.116]	[-0.343,0.244]	[-0.25,0.218]	[-0.07,0.046]	[-0.172,-0.071]	[-0.201,0.154]
Cognitive t	-0.15	-0.007	0.216	0.099	0.217	-0.16	-0.019	0.167	-0.064	0.151
	0.131	0.206	0.076	0.078	0.088	0.133	0.206	0.088	0.129	0.118
	[-0.369,0.05]	[-0.552,0.138]	[0.062,0.319]	[-0.019,0.224]	[0.072,0.359]	[-0.391,0.035]	[-0.577,0.121]	[-0.019,0.282]	[-0.321,0.083]	[-0.044,0.33]
Socio-emotional t	0.28	0.77	0.41	0.436	0.378	0.26	0.707	0.361	0.345	0.316
	0.159	0.216	0.077	0.067	0.092	0.148	0.229	0.089	0.075	0.109
	[0.118,0.609]	[0.304,0.983]	[0.332,0.568]	[0.332,0.543]	[0.242,0.548]	[0.106,0.538]	[0.208,0.912]	[0.258,0.538]	[0.211,0.454]	[0.148,0.495]
Health t	0.42	0.09	0.156	0.242	0.103	0.31	-0.044	0.05	0.144	0.09
	0.176	0.198	0.082	0.068	0.073	0.184	0.21	0.087	0.086	0.077
	[0.122,0.685]	[-0.106,0.556]	[0.051,0.315]	[0.16,0.38]	[-0.017,0.219]	[0.012,0.605]	[-0.287,0.425]	[-0.077,0.208]	[0.022,0.296]	[-0.069,0.201]
Parental cognition	0.04	-0.021	0.083	0.115	0.206	-0.01	-0.058	-0.084	-0.079	0.088
	0.111	0.129	0.068	0.056	0.089	0.152	0.146	0.095	0.094	0.149
	[-0.071,0.291]	[-0.054,0.37]	[-0.068,0.15]	[0.03,0.217]	[0.054,0.336]	[-0.323,0.145]	[-0.309,0.103]	[-0.309,-0.007]	[-0.231,0.06]	[-0.172,0.287]
Investments	0.41	0.167	0.135	0.108	0.095	0.6	0.415	0.506	0.654	0.354
	0.163	0.137	0.068	0.083	0.099	0.232	0.246	0.203	0.243	0.286
	[0.104,0.593]	[-0.052,0.386]	[0.043,0.264]	[-0.054,0.202]	[-0.049,0.257]	[0.259,0.944]	[0.223,0.989]	[0.256,0.92]	[0.368,1.136]	[-0.049,0.869]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).

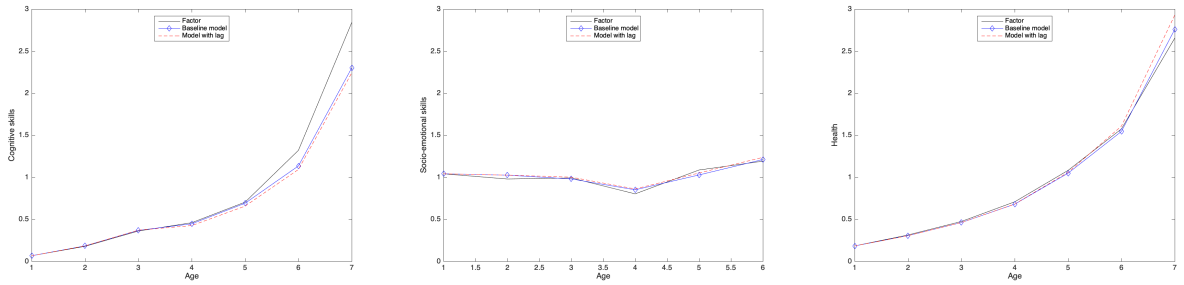
Table 7: Production functions for health

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.382	0.347	0.38	0.331	0.496	-0.06	0.402	0.368	0.353	0.35	0.497
	0.145	0.062	0.017	0.021	0.026	0.026	0.135	0.071	0.021	0.025	0.033	0.027
	[-0.14,0.319]	[0.252,0.449]	[0.323,0.379]	[0.338,0.407]	[0.305,0.391]	[0.455,0.538]	[-0.288,0.166]	[0.235,0.47]	[0.341,0.409]	[0.307,0.391]	[0.323,0.425]	[0.455,0.539]
Control function	-	-	-	-	-	-	-0.33	0.042	0.067	-0.09	0.132	0.033
							0.116	0.061	0.037	0.046	0.137	0.098
							[-0.473,-0.119]	[-0.102,0.106]	[0.015,0.128]	[-0.13,0.011]	[-0.012,0.433]	[-0.061,0.236]
AcioTu	-0.03	0.004	0.005	0.005	0.069	0.105	0	0.002	0.005	0.005	0.062	0.102
	0.061	0.042	0.005	0.007	0.032	0.028	0.05	0.041	0.005	0.007	0.032	0.03
	[-0.104,0.086]	[-0.069,0.059]	[-0.002,0.016]	[-0.009,0.014]	[0.004,0.105]	[0.05,0.137]	[-0.076,0.076]	[-0.068,0.06]	[-0.002,0.016]	[-0.01,0.014]	[-0.005,0.097]	[0.04,0.133]
Cognitive t	-0.02	0.061	0.002	-0.13	-0.1	-0.122	-0.03	0.062	0.008	-0.151	-0.075	-0.117
	0.063	0.1	0.022	0.033	0.038	0.033	0.062	0.098	0.022	0.036	0.049	0.036
	[-0.108,0.096]	[-0.043,0.268]	[-0.037,0.037]	[-0.191,-0.086]	[-0.152,-0.031]	[-0.177,-0.072]	[-0.108,0.093]	[-0.043,0.26]	[-0.03,0.044]	[-0.214,-0.093]	[-0.126,0.026]	[-0.171,-0.05]
Socio-emotional t	-0.01	0.069	0.026	0.042	-0.037	-0.035	-0.02	0.075	0.032	0.031	-0.014	-0.032
	0.098	0.052	0.011	0.014	0.036	0.027	0.087	0.05	0.014	0.016	0.049	0.028
	[-0.1,0.225]	[0.003,0.18]	[0.006,0.041]	[0.009,0.055]	[-0.088,0.028]	[-0.085,0.005]	[-0.1,0.187]	[0.006,0.169]	[0.012,0.057]	[-0.005,0.052]	[-0.07,0.079]	[-0.078,0.021]
Health t	0.77	0.892	0.95	1.107	1.196	1.237	0.71	0.904	0.963	1.095	1.202	1.239
	0.107	0.155	0.028	0.034	0.043	0.034	0.097	0.162	0.03	0.034	0.045	0.036
	[0.533,0.872]	[0.543,1.036]	[0.914,1.009]	[1.051,1.164]	[1.103,1.249]	[1.173,1.291]	[0.495,0.822]	[0.522,1.056]	[0.925,1.022]	[1.047,1.159]	[1.11,1.263]	[1.181,1.301]
Parental cognition	0	-0.006	-0.023	0.029	-0.017	-0.071	-0.03	-0.002	-0.002	0.005	0.027	-0.06
	0.071	0.029	0.015	0.02	0.035	0.032	0.058	0.031	0.018	0.023	0.063	0.049
	[-0.017,0.189]	[-0.026,0.062]	[-0.047,0.003]	[-0.013,0.052]	[-0.075,0.031]	[-0.123,-0.017]	[-0.062,0.095]	[-0.028,0.062]	[-0.03,0.03]	[-0.04,0.036]	[-0.049,0.16]	[-0.118,0.034]
Investments	0.27	-0.017	0.045	-0.048	-0.043	-0.009	0.37	-0.039	-0.002	0.02	-0.139	-0.031
	0.073	0.045	0.019	0.022	0.04	0.035	0.083	0.064	0.034	0.041	0.13	0.082
	[0.061,0.301]	[-0.061,0.082]	[0.01,0.069]	[-0.059,0.014]	[-0.102,0.023]	[-0.053,0.061]	[0.18,0.443]	[-0.117,0.099]	[-0.063,0.046]	[-0.038,0.101]	[-0.401,0.006]	[-0.196,0.061]

Notes: This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).

7 Windows of opportunity

In this Section, we use the estimated model to characterize the human development process in its various domains, focusing on aspects relevant for policy design. Specifically, we ask whether the development process exhibits *windows of opportunity* — periods in which interventions aimed at increasing parental investments are particularly effective and their effects sustainable. To this end, we simulate changes to the inputs of the production function and trace their long-run consequences through *impulse response functions*, an approach that naturally accounts for the dynamic interactions among the domains of development and the role of parental investment. Before turning to the simulations, however, we first assess the model's fit and present graphical evidence on the marginal productivity of parental investment at different ages.



(a) Cognitive skills (b) Socio-emotional skills (c) Health

Figure 3: Evolution of Skills Over Time (Model Estimates)

7.1 Model fit

We begin by using our production function estimates for the three domains of human development to assess how well the model can reproduce the developmental trajectories observed in the data. Figure 3 plots the evolution of cognitive skills, socioemotional skills, and health over time, measured as the average latent factor in each domain implied by the available measures and the estimated measurement system. Each panel also shows the trajectory predicted by our model, constructed by assigning each child the median initial conditions in the sample and simulating the estimated dynamics forward, holding investment fixed at its median value in each period. We report results from both the baseline model (Tables 5, 6, and 7) and the extended specification that includes additional self-productivity lags, with the latter estimates reported in the Appendix.

The figures show that, perhaps not surprisingly, our model fits remarkably well. However, for the cognitive factor, both the baseline model and the one with an additional lag in

cognition under-predict child development at 6 and, more severely, at 7. For the other two domains (socioemotional skills and health), the fit of both versions of the model is remarkably good.

7.2 Marginal Product of Investment

In Figure 4, we plot the marginal product of investment (MPI) by baseline skill level, implied by our production function estimates for cognitive skills, socio-emotional skills, and health (reported in Tables 5, 6, and 7).

The figure is organized as a 3×3 grid: the first row plots the MPI for cognitive skills, the second for socio-emotional skills, and the third for health. Each column shows how the marginal product varies with baseline level of cognitive skills (column 1), socio-emotional skills (column 2), and health (column 3), holding all other inputs at their sample median. Within each graph, we plot the MPI at three different ages, and note that the scale differs across graphs.

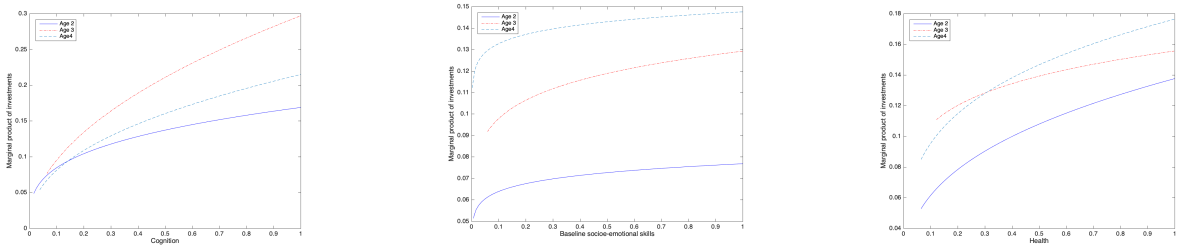
Because we estimate a Cobb-Douglas production function, the elasticity of each input is constant, but the marginal product itself is not: it depends on both the level of investment and the baseline level of skills. Holding investment at its median, the curvature of each MPI curve is governed by the estimated elasticity of the corresponding skill, so differences in curvature across graphs reflect differences in the estimated productivity of that skill.

We first note the complementarity between investment and the relevant skill domain (the diagonal graphs). (the diagonal graphs). This is particularly evident for cognition and socio-emotional skills, while for health it holds only at age 2. For cognition, the complementarity — reflected in the slope of the marginal productivity curves — becomes more pronounced at age 3. Investment also appears complementary with health and, to a lesser extent, with socio-emotional skills.

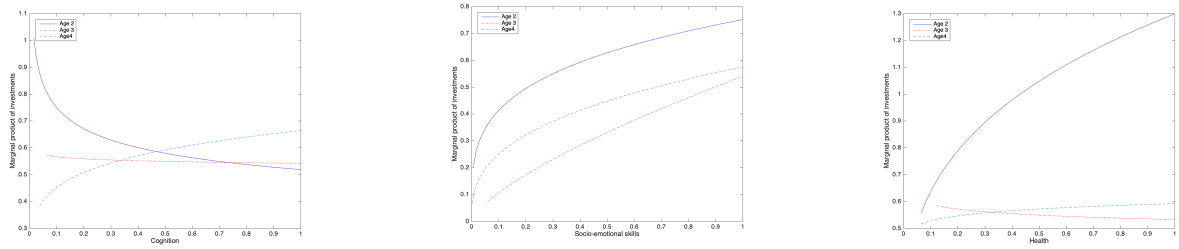
For socio-emotional skills, we find no strong complementarity between investment and cognitive development, nor, at ages 3 and 4, with health. Finally, for the health production function, we find no complementarity between investment and either cognition or socio-emotional skills.

7.3 Impulse response functions

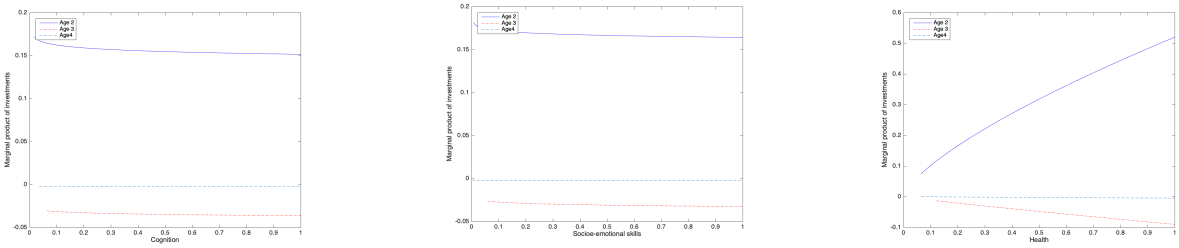
Figure 5 reports the impulse response functions (IRFs) for cognitive skills following exogenous innovations to parental investment implied by the estimated production functions. The IRFs are fully dynamic and incorporate the effects that a change in investment has, over time, on all domains of child development included in the model. In each panel, investments



(a) Cognitive skills



(b) Socio-emotional skills



(c) Health

Figure 4: Marginal Product of Investment

are increased by one standard deviation at a particular age (panels (a) to (d)). Panel (e) considers an increase in investment at ages 2, 3 and 4, while panel (f) considers an increase at ages 3 and 4. We compare the resulting evolution of cognitive skills with the counterfactual path that would have occurred had investments remained fixed at their median level in the data. At baseline, all other inputs are held at their median values. Each graph reports the predictions of both the baseline model and the model with an additional self-productivity lag, reported in the Appendix.

The IRFs reveal substantial heterogeneity in the timing and persistence of investment effects. The most striking result emerges from the comparison between panels (a) and (b). A one-standard-deviation increase in investment at age 2 raises cognitive skills by approximately 0.12 standard deviations on impact. This effect remains largely persistent through age 7, although the richer dynamic specification predicts a modest decline at ages 6 and 7. By contrast, an equally sized increase in investment at age 3 produces a similar contemporaneous effect but virtually no persistence beyond age 4. This difference arises because investment at age 2, unlike investment at age 3, has a substantial effect on contemporaneous health, which subsequently feeds into cognitive development through the cross-dimensional linkages embedded in the production function. Because this health effect is then propagated through the dynamic structure of the model, investments made at age 2 generate substantially larger long-run effects than investments made one year later.

The remaining panels further illustrate the importance of timing. Investment at age 4 also produces persistent effects on cognition, whereas investment at age 5 has virtually no effect. The comparison between panels (e) and (f) highlights not only the importance of intervening early, but also the importance of sustaining higher levels of investment over several periods. The cumulative effect of increasing investment at ages 2, 3 and 4 is substantially larger than that obtained when investment increases only at ages 3 and 4, and much larger than the isolated temporary shocks considered in panels (a) and (d).

These simulations are intended to illustrate the dynamic implications of the estimated production functions rather than to provide formal statistical inference. Confidence intervals are therefore omitted for brevity. Nevertheless, the results underscore the importance of modelling child development as a multidimensional and dynamic process. The interactions among cognitive, socio-emotional and health outcomes generate complex propagation mechanisms that have important implications for the identification of windows of opportunity for interventions and investments aimed at improving child development for disadvantaged children.

Figures 6 and 7 repeat the same exercise for socio-emotional skills and health, respectively. For socio-emotional skills, we again find that investment at age 2 generates the largest

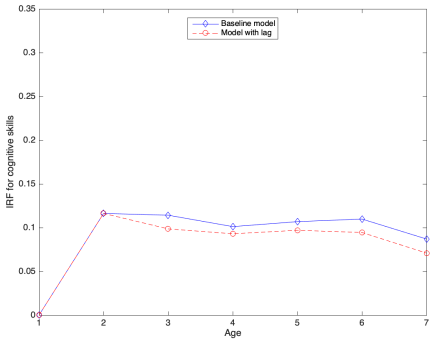
long-run effects. However, in contrast to cognition, the persistence of investment effects varies less across the timing of the innovation, and the predictions of the baseline model and of the specification with an additional self-productivity lag are remarkably similar. As in the case of cognition, the comparison between panels (e) and (f) shows that sustained increases in investment beginning at age 2 generate substantially larger gains than interventions that start later.

An interesting feature of the socio-emotional IRFs is the sizeable effect of investment at age 5. This stands in sharp contrast to the cognitive domain, for which investment innovations at age 5 have virtually no impact. Although we do not report the IRF corresponding to increases in investment from ages 2 through 5, the magnitude of the estimated coefficients and the patterns displayed in Figure 6 suggest that such an intervention would likely generate substantial improvements in socio-emotional development.

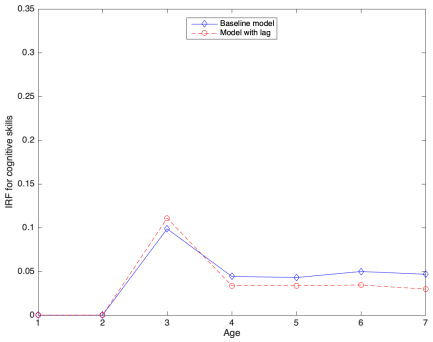
Finally, Figure 7 reports the IRFs for health. As before, we hold all other inputs fixed at their median values and report the predictions implied by both versions of the model. The results indicate that investment at age 2 has a persistent effect on health, whereas comparable investment shocks at later ages generate little persistence. Consistent with the cognitive IRFs, differences between the two specifications emerge only in the longer run and primarily for the age-2 innovation. As with cognition and socio-emotional skills, sustained investment across ages 2 through 4 generates substantially larger effects on health than isolated shocks at later ages (panels e and f).

These findings, consistent with the estimated coefficients of the health production function, are also important for understanding the importance of impacts generated by cross-domain effects. For instance, the strong response of health to investment at age 2 helps explain why early interventions generate persistent gains in cognition, and more generally why the timing of investments is a key determinant of their long-run effectiveness.

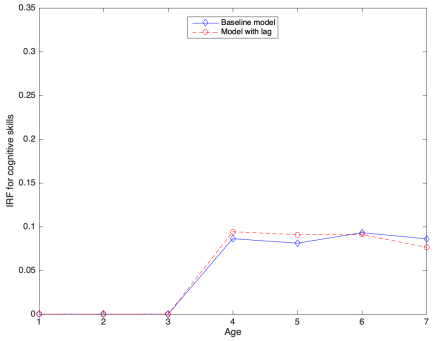
Taken together, the IRFs show that the effectiveness of parental investment depends not only on its magnitude but also on its timing and duration, and that these dynamics cannot be understood by looking at any single domain of development in isolation. Early investments in one domain — health, in particular — propagate through the cross-dimensional linkages in the production functions to generate persistent gains elsewhere, most notably in cognition. This transmission mechanism is what makes early investment disproportionately effective relative to equally sized investments made later, and it points to specific developmentally sensitive periods during which policy interventions are likely to be most effective.



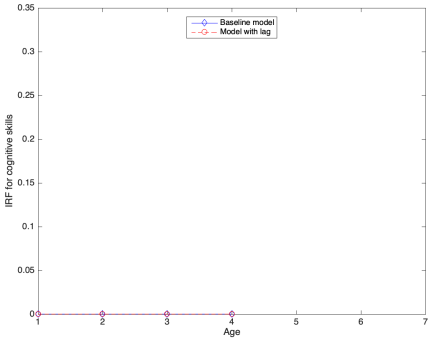
(a) Increase in investment at age 2



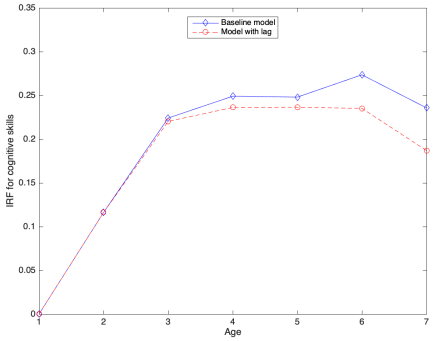
(b) Increase in investment at age 3



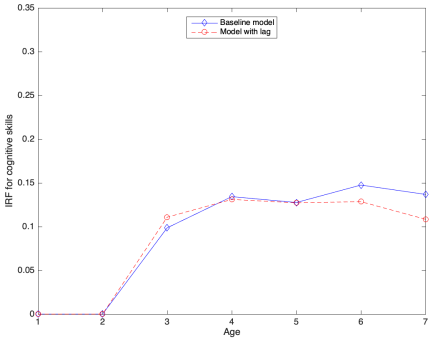
(c) Increase in investment at age 4



(d) Increase in investment at age 5



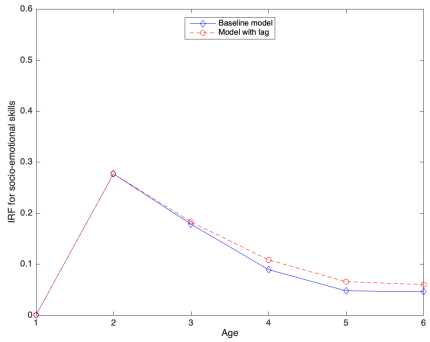
(e) Increase in investment at age 2, 3 and 4



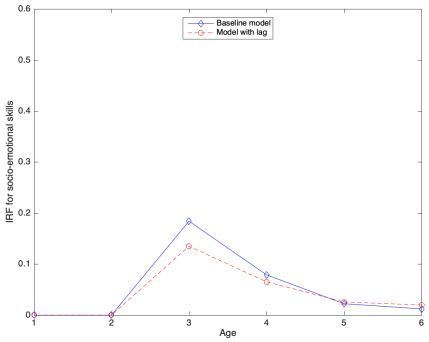
(f) Increase in investment at age 3 and 4

Figure 5: IRFs for cognitive skills

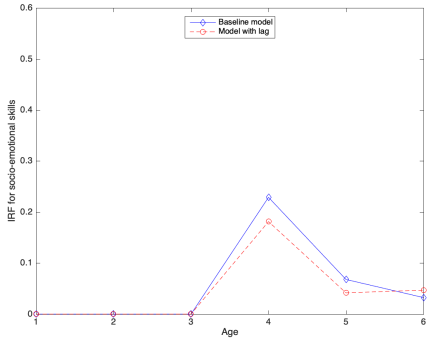
Notes:



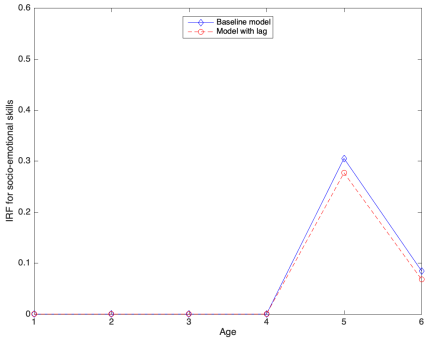
(a) Increase in investment at age 2



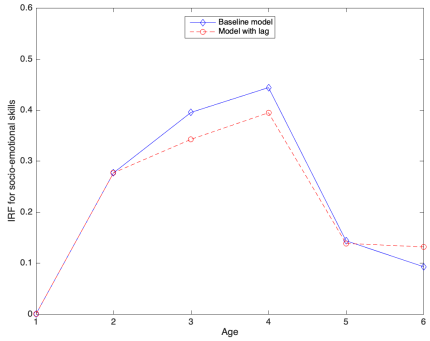
(b) Increase in investment at age 3



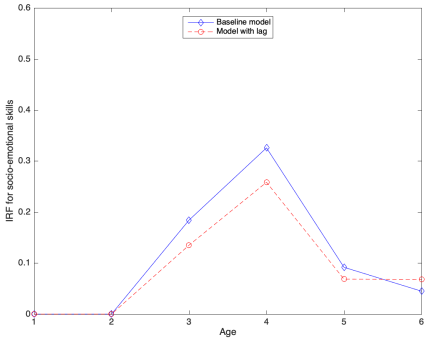
(c) Increase in investment at age 4



(d) Increase in investment at age 5



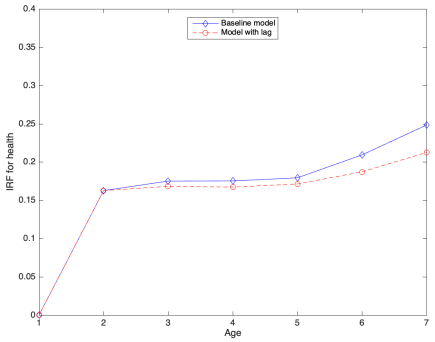
(e) Increase in investment at age 2, 3 and 4



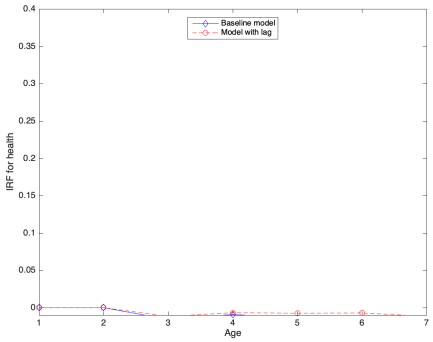
(f) Increase in investment at age 3 and 4

Figure 6: IRFs for socio-emotional skills

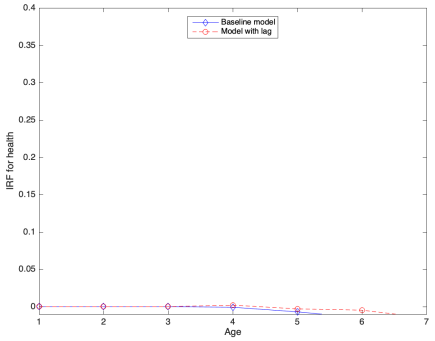
Notes:



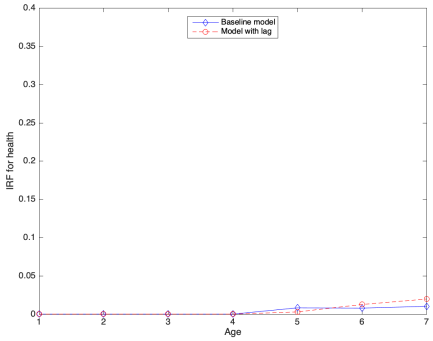
(a) Increase in investment at age 2



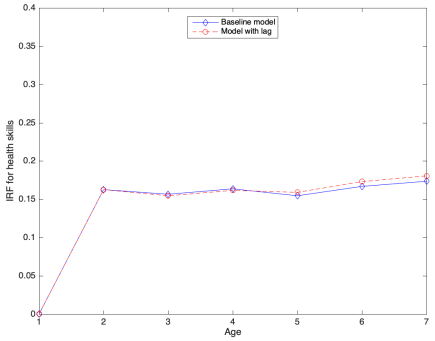
(b) Increase in investment at age 3



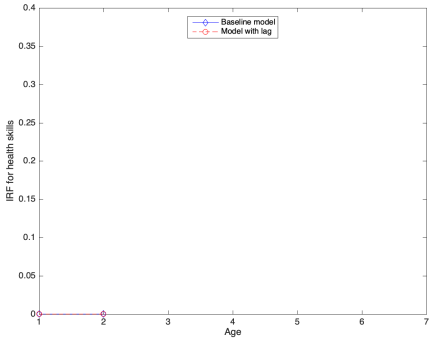
(c) Increase in investment at age 4



(d) Increase in investment at age 5



(e) Increase in investment at age 2, 3 and 4



(f) Increase in investment at age 3 and 4

Figure 7: IRFs for health

Notes:

8 Conclusions

In this paper, we study human development from ages 1 to 7 along three domains—cognition, socioemotional skills, and health—using a rich longitudinal dataset from Colombia that tracks, at relatively high frequency, a large sample of children from vulnerable families on the Caribbean coast. A distinctive feature of our analysis is the availability of repeated measures of child development at an annual frequency, which we use to characterize how the developmental process evolves with age.

A central methodological contribution is that we place the latent factors for each domain on an age-consistent scale, which is essential for dynamic comparisons. Some measures are not collected at all ages, and even when the same measure is observed over a broad age range, its mapping to the underlying developmental construct can change with age. Our method limits the distortions that would otherwise arise from these shifts and allows us to characterize developmental dynamics in a coherent metric over time.

Turning to the production function estimates, we find—consistent with the broader literature—that accounting for the endogeneity of parental investment is critical: ignoring it leads to a systematic underestimation of investment's productivity. At the same time, the productivity of investment varies sharply across both age and developmental domain. For health, investment matters primarily at age 2. For cognition, it is effective from ages 2 through 4, after which its influence wanes. Socioemotional skills exhibit a distinct pattern: investment remains consequential over a longer window and is particularly effective at age 5.

We also document meaningful changes over time in persistence within each domain. Most notably, cognitive development becomes substantially more persistent at age 5, whereas socioemotional skills display comparatively stable and lower persistence throughout the period we study. By contrast, we find little evidence of concurrent cross-domain interactions: current levels of one domain generally do not predict future levels of another. As discussed above, this does not rule out indirect complementarities operating through parental responses or other channels.

Taken together, the domain-specific patterns in investment productivity and persistence imply windows of opportunity that differ markedly across developmental domains, with direct implications for the design and timing of interventions. For example, our estimates suggest that very early investments in health can generate immediate improvements in that domain, which may in turn have longer-run consequences through health-related pathways for other dimensions of development, such as cognition.

Overall, our findings contribute to the literature on the technology of skill formation

by providing one of the most detailed characterizations of early child development available for a low- or middle-income country setting, with implications for both the optimal timing of policies and the assessment of their long-run sustainability. We complement the estimation of age-specific production functions with impulse response functions computed from the estimated model, which help interpret the estimates and distill their implications for developmental dynamics and, potentially, for policy design.

Our approach is not without limitations. First, while we model three domains of child development, socioemotional skills could plausibly be separated into internalizing skills (e.g., individual drive and the ability to focus) and externalizing skills (e.g., sociability and the ability to work in teams). Second, we do not distinguish between material and time investments, even though the literature suggests they may play different roles across domains. Some of these choices are constrained by the available measures and by the difficulty of constructing factors that are comparable across ages (and potentially across contexts). Richer measurement of child development and its determinants would substantially deepen our understanding of developmental processes.

Finally, because the data come from a single city and a disadvantaged population, we cannot assess how these relationships might differ in other environments or socioeconomic groups. The external validity of our evidence is therefore limited, underscoring the value of comparable longitudinal data across a wider range of contexts.

References

Agostinelli, Francesco and Matthew Wiswall, "Identification of Dynamic Latent Factor Models: The Implications of Re-Normalization in a Model of Child Development," 2016. NBER Working Paper No. 22441.

_ **and** _ , "Estimating the Technology of Children's Skill Formation," *Journal of Political Economy*, 2025, *133* (3), 846–887.

Anderson, T.W. and H. Rubin, "Statistical Inference in Factor Analysis," in "Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability," Vol. 5, Berkeley: University of California Press, 1956, pp. 111–150.

Attanasio, Orazio, Costas Meghir, and Emily Nix, "Human Capital Development and Parental Investment," *Review of Economic Studies*, 2020, *87* (1), 1–40.

_ , _ , _ , **and Francesca Salvati**, "Human capital growth and poverty: Evidence from Ethiopia and Peru," *Review of Economic Dynamics*, 2017, *11* (8).

_ , **Richard Blundell, Gabriella Conti, and Giacomo Mason**, "Inequality in noncognitive skills: a cross-cohort comparison," 2018. Unpublished.

_ , **Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina**, "Estimating the production function for human capital: results from a randomized controlled trial in Colombia," *American Economic Review*, 2020, *110* (1), 48–85.

Bayley, N., *Bayley Scales of Infant and Toddler Development*, 3rd ed., San Antonio, TX: Harcourt Assessment, 2006.

Bayley, Nancy, "Bayley scales of infant development," 2005.

Bernal, Raquel, Michele Giannola, and Milagros Nores, "The Effect of Center-Based Early Education on Disadvantaged Children's Developmental Trajectories: Experimental Evidence from Colombia," 2024. Available at SSRN: `https://ssrn.com/abstract=4191995`.

Christensen, Laurits R, Dale W Jorgenson, and Lawrence J Lau, "Transcendental logarithmic production frontiers," *The review of economics and statistics*, 1973, pp. 28–45.

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

Cunha, Filipe and James J. Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

DeBruin-Parecki, Andrea, *Early Literacy Skills Assessment User Guide.*, Ypsilanti, MI: High/Scope Press., 2005.

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

Heckman, James J., "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 2006, *312* (5782), 1900–1902.

Malaguzzi, Loris, "History, Ideas, and Basic Philosophy," in Carolyn Edwards, Lella Gandini, and George Forman, eds., *The Hundred Languages of Children: The Reggio Emilia Approach to Early Childhood Education*, Norwood, NJ: Ablex Publishing, 1993.

Muñoz-Sandoval, AF, RW Woodcock, KS McGrew, and N Mather, "Batería III Pruebas de aprovechamiento," *Itasca, IL: Riverside Publishing*, 2005.

Nelson, Charles A., "The Neurological Bases of Early Intervention," *Handbook of Early Childhood Intervention*, 2000, pp. 204–227.

Nores, M., R. Bernal, and S. Barnett, "Center-based care for infants and toddlers: the AEIOTU randomized trial," *Economics of Education Review*, 2019, *72*, 30 – 43.

Padilla, ER, DE Lugo, and LM Dunn, "Test de Vocabulario en Imágenes Peabody (TVIP)," *Circle Pines, MN: American Guidance Service (AGS), Inc.*, 1986.

Rubio-Codina, Marta, Caridad Araujo, Orazio Attanasio, Patricia Munoz, and Sally Grantham McGregor, "Concurrent Validity and Feasibility of Short Tests Currently Used to Measure Early Childhood Development in Large Scale Studies," *PLOS-ONE*, 2017, *25*, 234–259.

Shonkoff, Jack P. and Deborah A. Phillips, eds, *From Neurons to Neighborhoods: The Science of Early Childhood Development*, Washington, D.C.: National Academies Press, 2000.

Squires, Jane, Diane Bricker, and Elizabeth Twombly, *The ASQ: SE user's guide: For the Ages & Stages Questionnaires: Social-emotional.*, Paul H Brookes Publishing, 2002.

Appendix

In this appendix, we report the results we obtain estimating the production functions for the three domains of human development we consider (cognition, socio-emotional skills and health), introducing an additional lag of each of the domains considered. In particular, we modify Equation (1) to:

$$\theta_{t+1}^{k} = f_t(\boldsymbol{\theta_t}, \theta_{t-1}^{k}, X_t, \boldsymbol{Z_t}, \epsilon_{t+1}), k = Cog, SocEm, Health \tag{7}$$

where $\boldsymbol{\theta_t}$ is the vector $\{\theta_t^{cog}, \theta_t^{SocEm}, \theta_t^{Health}\}$, X_t is parental investment and $\boldsymbol{Z_t}$ is a vector of control variables. As with the previous tables, for each of the three dimensions of development we consider, we report results that treat parental investment as exogenous and results that deal with potential endogeneity using the same strategy as in the main text.

As can be seen from the tables, only in a few cases, the additional lag considered is statistically different from zero. And even then, these results do not affect the main implications discussed in the main text. This can also been seen in the simulations reported in the paper, where both sets of results were used.

Table 8: Production functions for cognitive skills with two self-productivity lags

	Cognitive t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	-0.19	0.202	-0.056	0.44	0.422	0.483	-0.29	0.097	-0.098	0.484	0.389	0.478
	0.16	0.102	0.043	0.047	0.073	0.109	0.154	0.101	0.047	0.062	0.083	0.11
	[-0.443,0.079]	[0.1,0.426]	[-0.121,0.012]	[0.353,0.505]	[0.342,0.575]	[0.527,0.876]	[-0.52,-0.037]	[-0.016,0.326]	[-0.199,-0.044]	[0.386,0.574]	[0.298,0.565]	[0.516,0.871]
Control function	-	-	-	-	-	-	-0.22	-0.303	-0.147	0.136	-0.135	-0.077
							0.145	0.107	0.093	0.112	0.157	0.122
							[-0.423,0.036]	[-0.459,-0.124]	[-0.361,-0.066]	[-0.011,0.327]	[-0.395,0.12]	[-0.235,0.145]
AcioTu	-0.13	0.034	0.056	0.068	0.006	0.046	-0.11	0.048	0.057	0.068	0.01	0.054
	0.081	0.057	0.016	0.018	0.046	0.043	0.078	0.051	0.016	0.018	0.047	0.045
	[-0.217,0.038]	[-0.049,0.127]	[0.028,0.08]	[0.025,0.089]	[-0.061,0.085]	[-0.009,0.125]	[-0.208,0.046]	[-0.044,0.127]	[0.028,0.081]	[0.026,0.087]	[-0.055,0.087]	[-0.013,0.132]
Cognitive t	0.3	0.469	0.319	1.022	1.281	1.321	0.3	0.435	0.3	1.044	1.271	1.315
	0.067	0.124	0.069	0.124	0.092	0.178	0.067	0.125	0.071	0.133	0.098	0.181
	[0.229,0.442]	[0.379,0.741]	[0.204,0.427]	[0.808,1.215]	[1.072,1.365]	[0.69,1.211]	[0.221,0.442]	[0.348,0.707]	[0.161,0.403]	[0.834,1.27]	[1.046,1.362]	[0.683,1.213]
Cognitive t-1	-	0.046	0.155	0	-0.248	-0.507	-	0.084	0.162	0.011	-0.275	-0.514
		0.076	0.066	0.073	0.115	0.199		0.075	0.065	0.073	0.117	0.202
		[-0.094,0.134]	[0.067,0.276]	[-0.12,0.101]	[-0.365,0.019]	[-0.377,0.205]		[-0.077,0.147]	[0.071,0.279]	[-0.111,0.121]	[-0.402,-0.014]	[-0.39,0.202]
Socio-emotional t	0.08	0.161	0.049	-0.06	0.053	0.082	0.08	0.123	0.035	-0.043	0.034	0.076
	0.103	0.077	0.023	0.031	0.044	0.042	0.101	0.075	0.027	0.035	0.052	0.043
	[-0.007,0.335]	[0.02,0.264]	[-0.002,0.074]	[-0.088,0.014]	[-0.061,0.085]	[0.006,0.143]	[-0.014,0.326]	[-0.017,0.216]	[-0.027,0.06]	[-0.072,0.039]	[-0.096,0.067]	[-0.012,0.14]
Health t	0.39	0.213	0.269	0.006	-0.03	-0.037	0.35	0.107	0.238	0.02	-0.032	-0.041
	0.12	0.1	0.047	0.051	0.045	0.035	0.116	0.095	0.054	0.053	0.046	0.036
	[0.134,0.512]	[0.034,0.355]	[0.178,0.339]	[-0.066,0.111]	[-0.099,0.044]	[-0.095,0.015]	[0.108,0.478]	[-0.017,0.271]	[0.132,0.309]	[-0.055,0.124]	[-0.105,0.044]	[-0.1,0.018]
Parental cognition	0.02	0.016	0.092	-0.011	0.006	0.12	0	-0.006	0.045	0.026	-0.035	0.092
	0.089	0.059	0.029	0.036	0.048	0.048	0.089	0.062	0.041	0.046	0.067	0.067
	[-0.017,0.233]	[-0.01,0.161]	[0.046,0.144]	[-0.054,0.064]	[-0.065,0.096]	[-0.009,0.149]	[-0.047,0.223]	[-0.092,0.087]	[-0.038,0.094]	[-0.023,0.117]	[-0.14,0.085]	[-0.048,0.171]
Investments	0.2	0.095	0.117	0.043	-0.06	0.02	0.27	0.258	0.221	-0.058	0.036	0.072
	0.085	0.068	0.038	0.044	0.053	0.046	0.12	0.106	0.086	0.103	0.143	0.099
	[-0.017,0.28]	[-0.038,0.176]	[0.066,0.194]	[-0.068,0.083]	[-0.131,0.049]	[-0.036,0.118]	[-0.032,0.376]	[0.088,0.401]	[0.147,0.416]	[-0.275,0.042]	[-0.165,0.297]	[-0.096,0.221]

Notes: This Table shows the estimates of the production function for cognitive skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).

Table 9: Production functions for socio-emotional skills with two self-productivity lags

	Socio-emotional t+1									
	Exogenous investments					Endogenous investments				
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 2	Age 3	Age 4	Age 5	Age 6
TFP	0.4	0.03	0.175	0.316	0.31	0.11	-0.079	0.06	0.107	0.272
	0.303	0.535	0.076	0.065	0.078	0.326	0.508	0.102	0.106	0.088
	[-0.143,0.812]	[-0.997,0.783]	[0.034,0.271]	[0.226,0.438]	[0.195,0.452]	[-0.483,0.597]	[-1.126,0.534]	[-0.133,0.177]	[-0.053,0.268]	[0.159,0.424]
Control function	-	-	-	-	-	-0.62	-0.258	-0.447	-0.683	-0.242
						0.279	0.243	0.25	0.268	0.347
						[-0.979,-0.097]	[-0.747,0.012]	[-0.935,-0.127]	[-1.228,-0.303]	[-0.749,0.369]
AeioTu	-0.12	0.147	0.035	-0.1	0.01	-0.08	0.158	0.012	-0.105	0.021
	0.181	0.15	0.038	0.033	0.097	0.173	0.148	0.039	0.034	0.102
	[-0.375,0.234]	[-0.222,0.24]	[-0.034,0.093]	[-0.158,-0.05]	[-0.18,0.125]	[-0.343,0.244]	[-0.217,0.271]	[-0.058,0.062]	[-0.16,-0.059]	[-0.191,0.142]
Cognitive t	-0.15	-0.041	0.212	0.093	0.17	-0.16	-0.04	0.173	-0.059	0.133
	0.131	0.272	0.077	0.081	0.094	0.133	0.266	0.088	0.124	0.116
	[-0.369,0.05]	[-0.721,0.182]	[0.051,0.31]	[-0.032,0.23]	[-0.003,0.312]	[-0.391,0.035]	[-0.716,0.138]	[-0.018,0.282]	[-0.286,0.1]	[-0.064,0.313]
Socio-emotional t	0.28	0.728	0.39	0.338	0.314	0.26	0.702	0.36	0.275	0.282
	0.159	0.211	0.085	0.076	0.103	0.148	0.224	0.09	0.078	0.111
	[0.118,0.609]	[0.245,0.918]	[0.293,0.557]	[0.193,0.448]	[0.149,0.484]	[0.106,0.538]	[0.173,0.887]	[0.248,0.534]	[0.116,0.369]	[0.099,0.46]
Socio-emotional t-1	-	0.087	0.171	0.114	0.164	-	0.066	0.075	0.088	0.139
		0.246	0.093	0.079	0.119		0.246	0.096	0.078	0.125
		[-0.265,0.61]	[0.034,0.3]	[0.054,0.283]	[0.013,0.393]		[-0.243,0.588]	[-0.081,0.215]	[0.015,0.242]	[-0.024,0.383]
Health t	0.42	0.078	0.118	0.263	0.093	0.31	0.005	0.05	0.166	0.086
	0.176	0.203	0.083	0.071	0.072	0.184	0.21	0.085	0.087	0.074
	[0.122,0.685]	[-0.121,0.555]	[0.003,0.268]	[0.162,0.397]	[-0.025,0.205]	[0.012,0.605]	[-0.246,0.468]	[-0.073,0.203]	[0.031,0.318]	[-0.043,0.196]
Parental cognition	0.04	-0.024	0.065	0.111	0.178	-0.01	-0.044	-0.066	-0.07	0.102
	0.111	0.112	0.066	0.055	0.088	0.152	0.122	0.097	0.09	0.144
	[-0.071,0.291]	[-0.097,0.249]	[-0.084,0.13]	[0.016,0.203]	[0.007,0.302]	[-0.323,0.145]	[-0.256,0.101]	[-0.28,0.018]	[-0.227,0.062]	[-0.148,0.312]
Investments	0.41	0.173	0.044	0.082	0.081	0.6	0.311	0.409	0.6	0.259
	0.163	0.15	0.071	0.089	0.095	0.232	0.22	0.22	0.247	0.28
	[0.104,0.593]	[-0.029,0.45]	[-0.042,0.199]	[-0.101,0.171]	[-0.067,0.245]	[0.259,0.944]	[0.087,0.843]	[0.204,0.895]	[0.222,1.072]	[-0.189,0.697]

Notes: This Table shows the estimates of the production function for socio-emotional skills. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).

Table 10: Production functions for health with two self-productivity lags

	Health t+1											
	Exogenous investments						Endogenous investments					
	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
TFP	0.1	0.394	0.298	0.256	0.076	0.376	-0.06	0.407	0.318	0.222	0.08	0.371
	0.145	0.065	0.026	0.072	0.095	0.059	0.135	0.068	0.028	0.079	0.097	0.062
	[-0.14,0.319]	[0.252,0.472]	[0.261,0.347]	[0.298,0.53]	[-0.022,0.28]	[0.306,0.486]	[-0.288,0.166]	[0.249,0.479]	[0.282,0.369]	[0.266,0.517]	[-0.003,0.3]	[0.301,0.498]
Control function	-	-	-	-	-	-	-0.33	0.034	0.06	-0.1	0.204	-0.032
							0.116	0.07	0.038	0.048	0.12	0.099
							[-0.473,-0.119]	[-0.112,0.101]	[0.002,0.12]	[-0.13,0.018]	[0.027,0.411]	[-0.125,0.184]
AcioTu	-0.03	0.004	0.009	0.001	0.028	0.074	0	0.003	0.009	0.001	0.014	0.076
	0.061	0.042	0.006	0.007	0.03	0.029	0.05	0.042	0.006	0.007	0.03	0.03
	[-0.104,0.086]	[-0.068,0.058]	[0.002,0.02]	[-0.008,0.014]	[-0.026,0.068]	[0.029,0.121]	[-0.076,0.076]	[-0.069,0.059]	[0.002,0.021]	[-0.008,0.014]	[-0.034,0.061]	[0.027,0.119]
Cognitive t	-0.02	0.05	0.012	-0.092	-0.032	-0.092	-0.03	0.054	0.017	-0.113	0.012	-0.096
	0.063	0.084	0.023	0.042	0.041	0.033	0.062	0.082	0.023	0.043	0.048	0.036
	[-0.108,0.096]	[-0.069,0.212]	[-0.031,0.047]	[-0.22,-0.077]	[-0.091,0.042]	[-0.157,-0.045]	[-0.108,0.093]	[-0.062,0.203]	[-0.025,0.054]	[-0.24,-0.094]	[-0.054,0.095]	[-0.158,-0.044]
Socio-emotional t	-0.01	0.067	0.025	0.038	-0.007	-0.01	-0.02	0.072	0.031	0.025	0.031	-0.012
	0.098	0.051	0.012	0.014	0.039	0.029	0.087	0.049	0.014	0.017	0.045	0.028
	[-0.1,0.225]	[0.002,0.168]	[0.005,0.041]	[0.008,0.056]	[-0.058,0.063]	[-0.068,0.032]	[-0.1,0.187]	[0.005,0.153]	[0.008,0.053]	[-0.004,0.054]	[-0.03,0.113]	[-0.063,0.028]
Health t	0.77	0.857	1.051	1.472	1.752	1.546	0.71	0.876	1.06	1.47	1.815	1.554
	0.107	0.233	0.066	0.215	0.208	0.152	0.097	0.248	0.067	0.217	0.206	0.16
	[0.533,0.872]	[0.381,1.119]	[0.955,1.175]	[0.639,1.327]	[1.278,1.97]	[1.26,1.744]	[0.495,0.822]	[0.334,1.132]	[0.965,1.186]	[0.654,1.335]	[1.323,2.006]	[1.226,1.743]
Health t-1	-	0.041	-0.117	-0.385	-0.698	-0.396	-	0.03	-0.113	-0.397	-0.766	-0.409
		0.144	0.06	0.223	0.252	0.176		0.149	0.061	0.228	0.249	0.191
		[-0.178,0.272]	[-0.214,-0.021]	[-0.219,0.491]	[-0.976,-0.155]	[-0.626,-0.066]		[-0.192,0.286]	[-0.208,-0.02]	[-0.229,0.488]	[-1.005,-0.193]	[-0.635,-0.018]
Parental cognition	0	-0.006	-0.019	0.034	0.012	-0.063	-0.03	-0.003	0	0.007	0.082	-0.074
	0.071	0.033	0.016	0.02	0.036	0.033	0.058	0.032	0.019	0.023	0.056	0.045
	[-0.017,0.189]	[-0.026,0.079]	[-0.041,0.011]	[-0.015,0.05]	[-0.053,0.056]	[-0.116,-0.008]	[-0.062,0.095]	[-0.028,0.063]	[-0.027,0.034]	[-0.038,0.036]	[-0.015,0.172]	[-0.119,0.014]
Investments	0.27	-0.009	0.047	-0.068	-0.027	0.015	0.37	-0.029	0.005	0.007	-0.174	0.036
	0.073	0.053	0.019	0.028	0.037	0.038	0.083	0.081	0.034	0.042	0.109	0.081
	[0.061,0.301]	[-0.069,0.095]	[0.01,0.073]	[-0.065,0.025]	[-0.089,0.028]	[-0.039,0.076]	[0.18,0.443]	[-0.119,0.141]	[-0.059,0.053]	[-0.035,0.107]	[-0.382,-0.037]	[-0.122,0.122]

Notes: This Table shows the estimates of the production function for health. The left panel considers investments as exogenous, while the right panel allows investments to be endogenous. 90% confidence intervals based on 200 bootstrap replications are reported in square brackets. Controls also include child age, gender, and the number of children in the household (see Table **??** for coefficients on these covariates).