Parental Investment and Child Development

Duc Thanh Nguyen*

June 10, 2024

Abstract

This paper studies the formation of cognitive skills, noncognitive skills and health. It analyzes the process by which current levels of cognitive skills, noncognitive skills and health depend on past cognitive and noncognitive abilities, past health, parental cognitive and noncognitive abilities and parental investments. I estimate dynamic, nested CES production function models of human capital with endogenous parental investments to examine dynamic complementarities and interactions among different inputs and factors in forming child human capital. I use a maximum likelihood approach by the minorization-maximization algorithm to estimate the joint distribution of latent factors, which are proxied by observable measures and dynamic CES production functions of human capital. My results show strong effects of parental investments on child cognitive skills, noncognitive skills and health and indicate that parental investments are driven by parental skills and resources. I find evidence that there are dynamic complementarities among the inputs in human capital production, implying that returns to investments are higher for children with better initial conditions. I also find evidence of high levels of self-productivity and the existence of cross-productivity from noncognitive skills and health to cognitive skills and from cognitive and noncognitive skills to health.

Keywords: Cognitive skills, noncognitive skills, health, dynamic factor analysis, dynamic skill accumulation, parental investments, child development, human capital.

*Department of Economics, Concordia University, e-mail: *ducthanhht@gmail.com*

1 Introduction

Developing human capital can offer a way for children to take advantage of new opportunities to improve their lives and contribute to sustainable economic growth and development. There is strong evidence showing that human capital is formed early in life and child development at early ages has long-lasting effects on adult social and economic outcomes (Knudsen, 2004; Knudsen et al., 2006; Cunha et al., 2006; Heckman et al., 2006; Urzúa, 2008; O'Neill, 1990). However, in developing countries, children face various risk factors and developmental deficits in every aspect of human capital development including cognitive, noncognitive skills and health that deter their development. Evidence shows that policies and interventions are effective in early childhood and for disadvantaged children (Knudsen et al., 2006; Doyle et al., 2009; Cunha et al., 2006; Cunha and Heckman, 2007; Heckman, 2008; Engle et al., 2007). There is also an increasing consensus that human capital is multidimensional with various components, including cognitive skills, noncognitive skills and health, and there are important dynamic complementarities and interactions among different components and factors. These dynamic complementarities and interactions, together with the fact that skills are malleable, give rise to potential early interventions and policies that can improve child development and thereby improve individual productivity. The effectiveness of such interventions and policies requires an understanding of the evolution of human capital throughout childhood: how its various components, including cognitive skills, noncognitive skills and health are formed and interacted, the importance of investments and the role of family background in driving child development and growth. However, our understanding of these mechanisms, roles and interactions is relatively limited to date.

In this study, I use high-quality data from the Vietnam Young Lives survey to estimate a dynamic production function model for the various dimensions of human capital with endogenous parental investments to examine dynamic complementarities and interactions among different inputs and factors in forming child human capital. I examine the process by which current stocks of cognitive skills, noncognitive skills and health depend on past cognitive and noncognitive skills, past health, parental cognitive and noncognitive skills, and parental investments. I use a maximum likelihood approach to estimate the joint distribution of the latent factors and dynamic CES production functions of human capital.

This research contributes to the existing literature in several ways. First, as far as I know, this is the first attempt to analyze the determinants and interactions of three important dimensions of human capital: cognitive skills, noncognitive skills and health. This research examines the dynamic production of these three key components of human capital that are likely to be fundamental determinants of children's productivity and future development. Second, it provides rare evidence in developing country settings. It utilizes high-quality longitudinal data from the Young Lives survey to identify the determinants and interactions of cognitive skills, noncognitive skills and health over two critical childhood development stages, aged 12 and 15. Third, the research uses a latent factor approach to identify the latent, unobserved factors instead of (noisy) proxy variables to correct measurement error problems, capture multiple skill dimensions more accurately and explore the endogeneity of investments.

2 Literature Review

Although research on skill foundation has been growing recently with a number of significant contributions (Cunha and Heckman, 2008; Cunha et al., 2010), the literature on skill foundation is still scarce. First, this is because it requires special longitudinal surveys that follow children throughout different periods of their life. The second reason arises from difficulties in directly measuring skills. Skill measures in survey data indirectly reflect true or latent cognitive and noncognitive skills.

The recent literature shows several important features of human capital formation. First, evidence shows that the development of human capital is dynamic, and the components of human capital - cognitive, noncognitive skills and health - are malleable and influenced by many external factors (Cunha and Heckman, 2007; Cunha et al., 2006; Attanasio, 2015). They are dynamically self-productive, i.e., the current stock of one human capital component begets the future stock of this own component and cross-productive, i.e., the current stock of one human capital component augments the development of another human capital component in the future. Cunha and Heckman (2008) show evidence of the self-productivity of skills, i.e., skills accumulated in one period foster the development of skills in future periods. Cunha and Heckman (2008) also find evidence of the cross-productivity of skills. In particular, their estimates show strong cross-productivity effects of noncognitive skills on cognitive skills, but the reverse seems weak. A second important feature of human capital development is dynamic complementarity, i.e. the productivity of investments in subsequent periods depends on skills, health and investments in previous periods. There is some evidence showing that parental investments play a key role in children's skill development (Doyle et al., 2009; Cunha and Heckman, 2008; Cunha and Heckman, 2007; Coneus et al., 2012). Evidence about skills' self-productivity, cross-productivity and dynamic complementarity show the importance of parental investments in children's early life (Cunha and Heckman, 2007).

While a series of research for developing countries show mixed results of self-productivity and cross-productivity, they find strong evidence of dynamic complementarity of human capital (Helmers and Patnam, 2011; Attanasio et al., 2017; Attanasio et al., 2020; Sánchez, 2017). For example, using the Young Lives data from India, Helmers and Patnam (2011) show that cognitive skills strongly affect both cognitive (self-productivity) and noncognitive skills (cross-productivity), but they find no evidence of self-productivity for noncognitive skills and cross-productivity effects from noncognitive to cognitive skills. This result differs from the results from Sánchez (2017), which favor the cross-productivity of noncognitive skills on cognitive skills. Attanasio et al. (2020)'s research on Mexico provides strong evidence of self-productivity, and they also find evidence of the cross-productivity effect from cognitive skills on noncognitive skills, but not vice versa. Consistent with Cunha and Heckman (2008), research in developing countries shows the importance of parental investments and dynamic complementarity in developing cognitive and noncognitive skills. Studying human capital development at early ages in developing countries is particularly important to boost policies, interventions and investments in children to minimize the loss of human potential given the evidence that they are exposed to various risk factors and face developmental deficits in every aspect of human capital development (Engle et al., 2007).

3 Model

3.1 Dynamics of Skill Formation

Literature has shown that human capital constituents (cognitive skills, noncognitive skills and health) are dynamic and influenced by many external factors. My framework builds on the dynamic factor models of Cunha et al. (2010), Cunha and Heckman (2007), Attanasio et al. (2017), Agostinelli and Wiswall (2016) and Aucejo and James (2021) to incorporate a variety of factors in the process of human capital production. In this framework, past skills and health produce future period skills and health; and investments can promote skill development and health and vice versa, the past stocks of skills and health can affect the next period's stock of skills and health indirectly by inducing investments in them, and the stocks of parental skills can affect their child's development. The current stocks of cognitive skills, noncognitive skills and health are determined by the past cognitive skills, noncognitive skills, health, parental investments and the parental stocks of cognitive skills and noncognitive skills. In particular, the stocks of cognitive and noncognitive skills and health of child i at time t (denoted by $\Theta^C_{i,t}$, $\Theta^{NC}_{i,t}$ and $\Theta^H_{i,t}$ respectively) are a function of the child's stock of cognitive skills, noncognitive skills and health at time t-1 (($\Theta^C_{i,t-1}$, $\Theta^{NC}_{i,t-1}$ and $\Theta^H_{i,t-1}$), the parental stocks of cognitive skills (P^C_i) and noncognitive skills (P^{NC}_i) and the investments made by the parent I_i.

$$\Theta^k_{i,t} = f(\Theta^C_{i,t-1}, \Theta^{NC}_{i,t-1}, \Theta^H_{i,t-1}, I_{i,t}, P^C_i, P^{NC}_i, X_{i,t}, A^k_t, v_{i,t}, \varepsilon^k_{i,t}) \tag{1}$$

$$k \in \{C, NC, H\}$$

I use a Constant Elasticity of Substitution (CES) production function. The CES production function has been recently used and it has been the most flexible functional form used in the human capital production literature. The CES functional form allows for a great level of flexibility in exploring substitutability between various inputs in the production function.

$$\begin{aligned} \Theta^k_{i,t} = \quad & \left[\gamma^k_{1t}(\Theta^C_{i,t-1})^{\rho^{tk}} + \gamma^k_{2t}(\Theta^{NC}_{i,t-1})^{\rho^{tk}} + \gamma^k_{3t}(\Theta^H_{i,t-1})^{\rho^{tk}} + \gamma^k_{4t}(I_{i,t})^{\rho^{tk}}\right. \\ & \left.\gamma^k_{5t}(P^C_i)^{\rho^{tk}} + \gamma^k_{6t}(P^{NC}_i)^{\rho^{tk}}\right]^{1/\rho^{tk}} e^{X'_{i,t}\delta^k_t + A^k_t + \mu^k v_{i,t} + \varepsilon^k_{i,t}} \end{aligned} \tag{2}$$

$$k \in \{C, NC, H\}$$

Where C, NC and H stand for cognitive skills, noncognitive skills and health respectively, $\gamma_{1t}^k + \gamma_{2t}^k + \gamma_{3t}^k + \gamma_{4t}^k + \gamma_{5t}^k + \gamma_{6t}^k = 1$.

In addition to the five different inputs mentioned above, I also include other components that contribute to the accumulation of human capital. $X_{i,t}$ are observable variables which include child background characteristics (gender of the child, the number of siblings in the family, residential region). $\varepsilon_{i,t}^k$ are normally distributed unobserved shocks. The term A_t^k represents total factor productivity (TFP). The production functions include the residual of a investment function, $v_{i,t}$, as a control function to control for endogenous investments that will be discussed in Sections 3.3 and 4 about parental investments and estimation.

3.2 Measurement System

The Young Lives survey contains rich data with multiple variables for human capital production functions. As discussed in Chapter 1, it is not efficient and feasible to use all of the available measures as separate variables in the production function. Furthermore, using measures observed in the data as a proxy for skills, investments and health suffers from measurement errors since all these measures provide imperfect proxies of latent skills, investments and health. These measures should only be considered as noisy, error-ridden proxies for latent unobserved skills, investment and health factors.

I use a latent factor model to extract the latent factors of interest from a large set of measures observed in the data and remove the measurement errors. The basic idea behind the factor approach is that one can relate measures observed in the data to unobserved, latent, underlying factors.

In this model, I estimate latent factors measuring child cognitive, noncognitive skills and health at time t ($\Theta_{i,t}^C$, $\Theta_{i,t}^{NC}$, $\Theta_{i,t}^H$), child cognitive, noncognitive skills and health at time t - 1 ($\Theta_{i,t-1}^C$, $\Theta_{i,t-1}^{NC}$, $\Theta_{i,t-1}^H$), parental cognitive and noncognitive skills (P_i^C, P_i^{NC}) and parental investments ($I_{i,t}$). Since the latent factors ($\Theta_{i,t}^C$, $\Theta_{i,t}^{NC}$, $\Theta_{i,t}^H$, $\Theta_{i,t-1}^C$, $\Theta_{i,t-1}^{NC}$, $\Theta_{i,t-1}^H$, P_i^C, P_i^{NC}, and $I_{i,t}$) are not directly measured, I use the factor model approach to extract these unobserved variables from a large set of observed data.

Since I estimate the log of the production function in Equation 2 and as required by the model that the factors are positive, I define the natural log of the factors as $\theta^k_{i,t} = ln(\Theta^k_{i,t})$, $\theta^k_{i,t-1} = ln(\Theta^k_{i,t-1})$, $\mathcal{P}^C_i = ln(P^C_i)$, $\mathcal{P}^{NC}_i = ln(P^{NC}_i)$ and $\mathcal{I}_{i,t} = ln(I_{i,t})$ so that latent factors only take positive values. With this definition, I assume that the observed measures proxy the natural log of the factors (Cunha et al., 2010; Agostinelli and Wiswall, 2016; Attanasio et al., 2017; Aucejo and James, 2021).

Let $T^k_{i,j,\tau}$ be the jth measure relating to latent factor k for individual i at time τ ($\tau \in \{t-1, t\}$. There are two types of the observed measures, continuous and binary measures. The continuous measures are described by:

$$T^k_{i,j,\tau} = \alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} \tag{3}$$

Where $\alpha^k_{j,\tau}$ are the intercepts, $\beta^k_{j,\tau}$ is the factor loadings on factor k for measure j at time τ. $u^k_{i,j,\tau}$ are measurement errors which are independent of the latent factors and mutually independent. $u^k_{i,j,\tau}$ reflect that the observed measures are imperfect proxies of the latent factors. The distributions of the error terms $u^k_{i,j,\tau}$, $f_{j,\tau}(.)$ are assumed to follow normal distributions with mean zero and variance $\sigma^2_{u_{j,\tau}}$, then he probability of the continuous measures is

$$f(T^k_{i,j,\tau}|\theta^k_{i,\tau}) = \frac{1}{\sqrt{2\sigma^2_{u_{j,\tau}}\pi}} exp\left(-\frac{(T^k_{i,j,\tau} - \alpha^k_{j,\tau} - \beta^k_{j,\tau}\theta^k_{i,\tau})^2}{2\sigma^2_{u_{j,\tau}}}\right) \tag{4}$$

The binary measures are described by:

$$T^k_{i,j,\tau} = \mathbb{1}[\alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} > 0] \tag{5}$$

Where $\mathbb{1}$ is an indicator function that equals one if $T^{*k}_{i,j,\tau} = \alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau} + u^k_{i,j,\tau} > 0$. $u^k_{i,j,\tau}$ are assumed to be logistically distributed, independent of the latent factors and mutually independent. Conditional on the unobservable factors, the probability of these binary measures is:

$$f(T^k_{i,j,\tau}|\theta^k_{i,\tau}) = \frac{exp(\alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau})^{T^k_{i,j,\tau}}}{1 + exp(\alpha^k_{j,\tau} + \beta^k_{j,\tau}\theta^k_{i,\tau})} \tag{6}$$

This equation maps the jth measure to latent, unobserved factor k. The assumption is that the observed measures are imperfect, error-ridden proxies for the underlying factors.

Specifically, the system of equations for continuous measures can be written as follows:

For child's skills and health:

$$\begin{aligned} T^{C}_{i,j,\tau} &= \alpha^{C}_{j,\tau} + \beta^{C}_{j,\tau}\theta^{C}_{i,\tau} + u^{C}_{i,j,\tau} \\ T^{NC}_{i,j,\tau} &= \alpha^{NC}_{j,\tau} + \beta^{NC}_{j,\tau}\theta^{NC}_{i,\tau} + u^{NC}_{i,j,\tau} \\ T^{H}_{i,j,\tau} &= \alpha^{H}_{j,\tau} + \beta^{H}_{j,\tau}\theta^{H}_{i,\tau} + u^{H}_{i,j,\tau} \end{aligned} \tag{7}$$

Parental skills and parental investments follow the same structure.

For parental skills:

$$\begin{aligned} T^{PC}_{i,j} &= \alpha^{PC}_{j} + \beta^{PC}_{j}\mathcal{P}^{C}_{i} + u^{PC}_{i,j} \\ T^{PNC}_{i,j} &= \alpha^{PNC}_{j} + \beta^{PNC}_{j}\mathcal{P}^{NC}_{i} + u^{PNC}_{i,j} \end{aligned} \tag{8}$$

For parental investments:

$$T^{I}_{i,j,\tau} = \alpha^{I}_{j,\tau} + \beta^{I}_{j,\tau}\mathcal{I}_{i,\tau} + u^{I}_{i,j,\tau} \tag{9}$$

The system of equations for binary measures can be specified as follows:

$$\begin{aligned} T^{C}_{i,j,\tau} &= \mathbb{1}[\alpha^{C}_{j,\tau} + \beta^{C}_{j,\tau}\theta^{C}_{i,\tau} + u^{C}_{i,j,\tau} > 0] \\ T^{NC}_{i,j,\tau} &= \mathbb{1}[\alpha^{NC}_{j,\tau} + \beta^{NC}_{j,\tau}\theta^{NC}_{i,\tau} + u^{NC}_{i,j,\tau} > 0] \\ T^{H}_{i,j,\tau} &= \mathbb{1}[\alpha^{H}_{j,\tau} + \beta^{H}_{j,\tau}\theta^{H}_{i,\tau} + u^{H}_{i,j,\tau} > 0] \\ T^{PC}_{i,j} &= \mathbb{1}[\alpha^{PC}_{j} + \beta^{PC}_{j}\mathcal{P}^{C}_{i} + u^{PC}_{i,j} > 0] \\ T^{PNC}_{i,j} &= \mathbb{1}[\alpha^{PNC}_{j} + \beta^{PNC}_{j}\mathcal{P}^{NC}_{i} + u^{PNC}_{i,j} > 0] \\ T^{I}_{i,j,\tau} &= \mathbb{1}[\alpha^{I}_{j,\tau} + \beta^{I}_{j,\tau}\mathcal{I}_{i,\tau} + u^{I}_{i,j,\tau} > 0] \end{aligned} \tag{10}$$

Since the underlying/latent factors are unobserved, for identification, we need to normalize one of the loadings for each factor to one and one of the intercepts for each factor to zero. Since a variety of measures that may change from age to age are used, each factor is normalized on the same measure at every age to make the comparisons over time consistent. Child cognitive skills are always normalized on the Peabody Picture Vocabulary Test (PPVT), child noncognitive skills are normalized on self-esteem and child health is normalized on height. Parental cognitive skills are normalized on the mother's years of education, Parental noncognitive skills are normalized on parental self-esteem and parental investments are normalized on expenditure on the Young Lives child. Another condition to identify factors is that the number of observable measures $L >= 2k + 1$, where L is the number of measures and k is the number of factors (Cunha et al., 2010). This condition is satisfied since there are 27 measures for 9 factors in the model.

Let $\theta_i = [\theta^k_{i,\tau}]' = [\theta^C_{i,\tau}, \theta^{NC}_{i,\tau}, \theta^H_{i,\tau}, \mathcal{P}^C_i, \mathcal{P}^{NC}_i, \mathcal{I}_{i,\tau}]'$ be a complete vector of realized factors of individual i, from Equations 4 and 6, the likelihood of all the observed measures conditional on θ_i is

$$L(T_i|\theta_i) = \prod_{k=1}^{K}\prod_{j=1}^{J}\prod_{\tau=1}^{T} f(T^k_{i,j,\tau}|\theta_i) \tag{11}$$

The distributions of the log factors, $f(\theta)$, are assumed to be jointly distributed as a mixture of two normals. The distributions of the latent factors may follow many forms and the assumption of the factor distributions is important and must be flexible enough to capture data and models. The assumption of the mixture of normal distributions of the factors is essential in this model. First, fewer restrictions are imposed on the distributions and the distributions are flexible enough to capture data. Second, the production function functions are non-linear, so the distributions need to be general and flexible enough to be consistent with the model (Attanasio et al., 2017). The probability density function of the factor is then $f(\theta) = \sum_{c=1}^{2} \tau_c f(\theta|\mu_c, \Omega_c)$, where μ_c, Ω_c and τ_c are the mean, covariance and the mixture probability of the two normals.

Let Ψ be all the parameters of the model, $\Psi = \{\alpha, \beta, \sigma, \tau_c, \mu_c, \Omega_c\}$. The log-likelihood function is:

$$\begin{aligned}\mathcal{L}(\Psi) &= \sum_{i=1}^{N} lnL(T_i) \\ &= \sum_{i=1}^{N} ln\left(\int L(T_i|\theta_i)dF(\theta)\right) \\ &= \sum_{i=1}^{N} ln\left(\int L(T_i|\theta_i)f(\theta)d\theta\right)\end{aligned} \tag{12}$$

Given the unobservable nature of the factors, the log-likelihood function is constructed by integrating over the distributions of the unobservable factors.

3.3 Parental Investments

Parental investments reflect parents' choices and they depend on parents' objectives, resources and how effective the investments in their children are. Investments are endogenously determined by parental resources, expectations regarding the returns to investments in their children, and the parent's levels of cognitive and noncognitive skills. Parents make investment choices taking into account the child's stocks of cognitive and noncognitive skills and health since returns to investments may depend on their child's stocks of human capital, in particular, if the child's stocks of human capital and investments are complementary. Investments may depend on the parent's levels of cognitive and noncognitive skills because parents with higher levels of human capital may be better aware of the value of investments and may have higher lifetime resources.

Parental investments are an input in the production function and reflect parents' choices considering the evolution of the child's human capital. Parents react to their child's human capital when they choose their investments in their children. Therefore, parental investments could be endogenous. Parental investments could be correlated with unobserved shocks or omitted inputs that are relevant for child human capital accumulation.

To deal with the endogenous nature of parental investments, I use the household wealth index, economic shocks and regional prices as instruments and use a control function approach inspired by Attanasio et al. (2017). These instruments are valid provided that wealth index, economic shocks and regional prices affect cognitive and noncognitive skills and health only through their impacts on parental investments.

In particular, the parental investment function is specified as follows:

$$\begin{aligned} lnI_{i,t} = \quad & \alpha_{1,t} + \alpha_{2,t} ln\Theta^{C}_{i,t-1} + \alpha_{3,t} ln\Theta^{NC}_{i,t-1} + \alpha_{4,t} ln\Theta^{H}_{i,t-1} + \alpha_{5,t} lnP^{C}_{i} + \alpha_{6,t} lnP^{NC}_{i} \\ & + \alpha_{7,t} X_{i,t} + \alpha_{8,t} Z_{i,t} + v_{i,t} \end{aligned} \tag{13}$$

Where $X_{i,t}$ includes child gender, urban/rural residence and the number of siblings. $Z_{i,t}$

a vector of the instrumental variables that determine the parental investment choices and are not included in the production function. $Z_{i,t}$ are the log of wealth index reflecting parental resources, household economic shocks (Shock in input prices[1], drought, flood, crop failure and illness of household members) and the log of regional prices. $v_{i,t}$ is an error term.

Because data on family income are not available in the survey, the wealth index is used to proxy for parental resources. Parental resources and prices included in the model reflect budget constraints. This model can be considered as an approximation to a dynamic model of household choice and parental investments with liquidity constraints in which parents make investment choices to maximize a welfare function with arguments of human capital and consumption, subject to a budget constraint and the production functions (Del Boca et al., 2013; Attanasio et al., 2017).

4 Estimation

The model estimation consists of two steps. In the first step, I estimate the measurement system to recover the parameters $\beta^C_{j,\tau}$, $\beta^{NC}_{j,\tau}$, $\beta^H_{j,\tau}$, β^{PC}_j, β^{PNC}_j, $\beta^I_{j,t}$, α^C_j, α^{NC}_k, α^{PC}_j, α^{PNC}_j, $\alpha^I_{j,t}$ and the latent factor distributions by maximum likelihood estimation (MLE). In particular, I estimate the log-likelihood function 12 using the minorization-maximization algorithm that is presented in Appendix B. In the second step, I use the estimated parameters of the factor distributions from the first step to take individual-specific draws and use these draws as observable data to estimate investment and production functions.

The parental investment function takes the form of Equation 13:

$$\begin{aligned} lnI_{i,t} = \; & \alpha_{1,t} + \alpha_{2,t} ln\Theta^C_{i,t-1} + \alpha_{3,t} ln\Theta^{NC}_{i,t-1} + \alpha_{4,t} ln\Theta^H_{i,t-1} + \alpha_{5,t} lnP^C_i + \alpha_{6,t} lnP^{NC}_i \\ & + \alpha_{7,t} X_{i,t} + \alpha_{8,t} Z_{i,t} + v_{i,t} \end{aligned} \tag{14}$$

The $v_{i,t}$ is the residual of the investment function as a control function. In this specification, household wealth index, economic shocks and regional prices are included in the

[1] Shock in input prices refers to a large increase in the prices of inputs such as fertilizers, plant seeds or machinery and equipment for agricultural production.

investment function but not in the production functions as follows:

$$ln\Theta_{i,t}^{k} = ln(g(\Theta_{i,t-1}^{C}, \Theta_{i,t-1}^{NC}, \Theta_{i,t-1}^{H}, I_{i,t}, P_{i}^{C}, P_{i}^{NC})) + X_{i,t}'\delta_{t}^{k} + A_{t}^{k} + \mu^{k}v_{i,t} + \varepsilon_{i,t}^{k} \tag{15}$$

Where $g(.)$ is the CES production function indicated earlier.

5 Data and Variables

I use the data for the Older Cohort from the Young Lives survey in Vietnam that follows 1,000 children from the age of 8 to age 22. It provides a rich data set on individual, family and community characteristics, health and cognitive and noncognitive skills. This research uses data from Round 2 (at age 12) and Round 3 (at age 15). This is because I want to investigate the impact of investments on child skills and health during adolescence and these two rounds contain cognitive, noncognitive skill and health measures needed for the research.

The household survey contains information on an extensive set of socio-economic and demographic characteristics, alongside a wealth of information around parenting, parental characteristics, and maternal skills, including mothers' years of education, verbal ability, IQ, depressive symptoms, and knowledge of child development. Table 1 presents descriptive statistics on the general characteristics of the sample. The sample includes 961 children. There is a balance between boys and girls. Around 80% of the child live in rural areas.[2] On average, the number of children in the household is around 1.

[2] Young Lives in Vietnam uses a pro-poor sampling strategy but represents the diversity of children in the country (Nguyen, 2008). This sampling design allows for studying the human capital development of young people in relatively low-resource settings.

Table 1: Key Descriptive Statistics

	Age 12, Round 2	Age 15, Round 3
Female	0.504	0.504
	(0.500)	(0.500)
Urban	0.198	0.199
	(0.398)	(0.399)
Number of siblings aged 0-18	1.347	0.965
	(0.982)	(0.947)
Wealth index	0.537	0.624
	(0.178)	(0.181)
Observations	961	961

Note: Standard deviations in parentheses.

Table 2 shows how the jth observed measure is mapped to the kth latent/underlying factor. The first loading of each factor is normalised to unity and thus the scale of the latent factors is defined by these measures. The next subsections 5.1, 5.2 and 5.3 describe the variables and corresponding latent factors used in the measurement system.

5.1 Children's Measures: Cognitive Skills, Noncognitive Skills and Health

Children's cognitive skill indicators for Round 3 are measured by the test scores in Peabody Picture Vocabulary Test (PPVT), mathematics test (math test), and reading comprehension test (Cloze test), and children's cognitive skill measures for Round 2 are measured by the PPVT score, math test score, and children's reading and writing levels.

Math test: The math test was administered in Rounds 2 and 3. It includes 29 items on addition, subtraction, multiplication, division, problem-solving, measurement, data interpretation, and basic geometry.

PPVT: The PPVT is a widely-used test of receptive vocabulary. It uses a stimulus word and a set of accompanying pictures to test receptive vocabulary. It has been used extensively to demonstrate the correlation between PPVT scores and cognitive and intellectual ability (Walker et al., 2005). The 204−item PPVT−III was used in Vietnam. Young Lives

Table 2: Observed Variables in the Young Lives Surveys and Corresponding Latent Factors

Latent factors		**Observed variables**
Child's cognitive skills - Round 2	θ_2^C	1. PPVT test 2. Math Test 3. Reading level 4. Writing level
Child's cognitive skills - Round 3	θ_3^C	1. PPVT test 2. Math Test 3. Cloze
Child's noncognitive skills - Round 2 and Round 3	θ_2^{NC}, θ_3^{NC}	1. Self-esteem score 2. Self-efficacy score 3. Self-respect and inclusion score
Child's health - Round 2 and Round 3	θ_2^H, θ_3^H	1. Child height for age z-score 2. Child weight 3. How is child health?
Parental cognitive skills	$\mathcal{P}^C$	1. Mother's years of education 2. Father's years of education
Parental noncognitive skills	$\mathcal{P}^{NC}$	1. Self-esteem score 2. Self-efficacy score 3. self-respect and inclusion score
Parental Investments	$\mathcal{I}_3$	1. Expenditure on the Young Lives child 2. Number of hours studying outside school as a proxy for the time that parents dedicate to the child 3. Quality of relationship between child and parents

researchers in each country followed a standard process for adaptation and standardization of the PPVT.

Cloze: The Cloze test was developed to measure verbal skills and reading comprehension. The test includes 24 items that increase in difficulty. Each item consists of a sentence or short paragraph that lack one or more words, children were asked to identify a word that completed the meaning of the sentence or paragraph. A thorough analysis of psychometric characteristics was examined to establish the reliability and validity of all these tests (Crookston et al., 2014).

I use three indicators designed to access dimensions of self-esteem, self-efficacy and self-respect and inclusion to measure children's noncognitive skills.[3]

Self-esteem: Self-esteem measures aspects related to pride and it builds on the Rosenberg scale (Rosenberg, 1965).

Self-efficacy: The self-efficacy scale measures aspects related to agency and it builds on the Rotter scale (Rotter, 1966).

Self-respect and inclusion: focuses on the social component of self-esteem (Dercon and Krishnan, 2009). The statements used to measure self-respect revolve around the concepts of pride and the sense of inclusion.

Self-esteem and self-efficacy have been extensively studied and widely used. They have been validated and proved reliable in psychological and economic literature. Self-esteem and self-efficacy are the most common noncognitive skill variables used in empirical studies (Glewwe et al., 2017). Self-respect and inclusion are related to the self-esteem measure but focus on the social and psychosocial aspects of inclusion. The single measures of self-esteem, self-efficacy and self-respect and inclusion are set on a Likert scale ranging from "strongly disagree" to "strongly agree". Children were read statements and asked whether they strongly disagreed, disagreed, more or less, agreed or strongly agreed with the statements. Negative statements are recoded to reflect positive statements.

Child health measures used in this study include height for age z-score, weight and self-rated health status. The rationale for using these measures is because height for age z-score may reflect the information of longer-term health and nutrition status and it is calculated

[3] The items/statements used to assess these scales are described in Appendix A.

Table 3: Key Descriptive Statistics - Child Measures

	Age 12, Round 2	Age 15, Round 3
Cognitive skill measures:		
PPVT test	137.900	166.471
	(25.336)	(27.702)
Math test	7.465	17.810
	(1.857)	(7.607)
Cloze test	-	17.934
		(4.979)
Reading level	0.972	-
	(0.166)	
Writing level	0.942	-
	(0.233)	
Noncognitive skill measures:		
Self-Esteem, raw score	3.446	3.850
	(0.443)	(0.569)
Self-Efficacy, raw score	3.394	4.168
	(0.367)	(0.504)
Self-respect and inclusion, raw score	3.538	3.778
	(0.397)	(0.421)
Health measures:		
Height-for-age z-score	-1.457	-1.420
	(1.081)	(0.909)
Child's weight (kg)	32.990	44.268
	(6.570)	(7.127)
How is child health?	0.715	0.503
	(0.452)	(0.500)
Observations	961	961

Note: Standard deviations in parentheses.

according to the World Health Organization (WHO) standards, while weight likely captures short-term health status. Height for age, weight and self-rated health are used to capture nutrition and health status in several studies using the Young Lives survey data (Helmers and Patnam, 2011; Attanasio et al., 2017; Sánchez, 2017).

5.2 Parental Cognitive Skills and Noncognitive Skills

In the models, parental cognitive skills and noncognitive skills are used to control for parental background. I use maternal education and paternal education to measure parental cognitive skills. Parental noncognitive skills are measured by three indicators designed to access dimensions of self-esteem, self-efficacy and self-respect and inclusion.[4] Parental cognitive and noncognitive skills are measured and treated as fixed in Round 2. The single measures of self-esteem, self-efficacy and self-respect and inclusion are set on a four-point Likert scale ranging from "strongly disagree" to "strongly agree". Parents were read statements and asked whether they strongly disagreed, disagreed, agreed or strongly agreed with the statements. Negative statements are recoded to reflect positive statements.

5.3 Parental Investments

I use three variables that measure parental resources devoted to the child in terms of money, time and the quality of the relationship between the parent and the child to extract the latent factor of parental investments. The first variable measures material investments that include expenditure on education, clothing, shoes, and books specifically devoted to the child. The second variable is the average number of hours per day the child studied outside school as a proxy for the time that parents dedicated to the child. The last variable measures the quality of the relationship between the child and the parent.[5]

[4] The construction of these indicators is described in Appendix A.
[5] This scale is constructed using the items listed in Appendix A.

Table 4: Key Descriptive Statistics - Parental Skill Measures

Cognitive skill measures:	
Mother's years of education	6.014
	(3.707)
Father's years of education	6.829
	(3.680)
Noncognitive skill measures:	
Self-Esteem, raw score	3.536
	(0.445)
Self-Efficacy, raw score	3.587
	(0.548)
Self-respect and inclusion, raw score	3.653
	(0.445)
Observations	961

Note: Standard deviations in parentheses.

Table 5: Descriptive Statistics: Parental Investments

	Round 2, age 12	Round 3, age 15
Expenditure on the Young Lives child	851.094	2559.822
	(985.913)	(3005.632)
Study hours outside school	2.890	3.066
	(1.597)	(2.117)
Quality of relationship, raw score	3.367	2.663
	(0.405)	(0.420)
Observations	961	961

Note: Standard deviations in parentheses.

6 Results

6.1 Measurement System

The measurement system relies on observed measures to identify the latent factors and it is used to assess how factors load on each of the measures.

To assess the information content contained in each measure from factors and measurement errors, I calculate the contribution of latent factors and measurement errors in explaining the variance of the observed measures.

$$P_j^{\theta_\tau^k} = \frac{(\beta_{j,t}^k)^2 var(\theta_{i,\tau}^k)}{(\beta_{j,\tau}^k)^2 var(\theta_{i,\tau}^k) + var(u_{i,j,\tau}^k)} \tag{16}$$

$$P_j^{u_{j,\tau}^k} = \frac{var(u_{i,j,\tau}^k)}{(\beta_{j,\tau}^k)^2 var(\theta_{i,\tau}^k) + var(u_{i,j,\tau}^k)} \tag{17}$$

Where $P_j^{\theta_\tau^k}$ is the proportion of the variance of the jth observed measure explained by latent factor k at time τ or signal and $P_j^{u_{j,\tau}^k}$ is the variance of the measure explained by the measurement error or noise.

Table 6 shows the measures assigned to each factor and the estimates of factor loadings onto the log of the factors. It also reports the fraction of the variance of each measure explained by each factor (signal) and by the measurement error (noise). I find that, for the most part, the measures proposed in Section 5 contain a substantial amount of information for each factor. The cognitive skill factors in both rounds account for an important fraction of the variance of each observed measure - from 45.5% to 69.8%. The measures on noncognitive skills are also very informative, from 15.9% to 68.7% of the variance of the noncognitive measures are accounted for by signal. Similarly, the latent factors of parental cognitive and noncognitive skills and parental investments explain an important proportion of the variance of the related observed measures, from 16.1% to 99.5% with the exceptions of parental self-efficacy and the parental quality of relationship with the child which are close to zero. The factor for health explains an considerable share of the variance of the health indicators with the signals exceeding 50%. Although the factors explain an essential proportion of the variance of the observed measures, these proportions are far from 100%. This shows that the observed measures capture the latent/true factors with significant measurement errors and demonstrates the importance of the latent factors in assessing human capital accumulation.

Table 6: Measurement System[6]

Latent factors and measures	Data type	Loading	AME	Signal	Noise
Child's cognitive skills - Round 3, age 15					
1. PPVT test	Continuous	1	0.765	0.572	0.428
2. Math test	Continuous	1.071	0.819	0.649	0.351
3. Cloze	Continuous	0.899	0.687	0.455	0.545
Child's cognitive skills - Round 2, age 12					
1. PPVT test	Continuous	1	0.834	0.698	0.302
2. Math Test	Continuous	1.020	0.851	0.603	0.397
3. Reading	Binary	1.687	0.018	–	–
4. Writing	Binary	1.530	0.035	–	–
Child's noncognitive skills - Round 3, age 15					
1. Self-esteem score	Continuous	1	0.381	0.358	0.642
2. Self-efficacy score	Continuous	0.759	0.289	0.308	0.692
3. Self-respect and inclusion score	Continuous	1.200	0.457	0.687	0.313
Child's noncognitive skills - Round 2, age 12					
1. Self-esteem score	Continuous	1	0.435	0.561	0.439
2. Self-efficacy score	Continuous	0.539	0.234	0.159	0.841
3. Self-respect and inclusion score	Continuous	0.773	0.336	0.327	0.673
Child's health - Round 3, age 15					
1. Height for age z-score	Continuous	1	0.623	0.391	0.609
2. Weight	Continuous	1.513	0.942	0.891	0.109
3. How is child health?	Binary	0.636	0.094	–	–

Continued on next page

Table 6 – *Continued from previous page*

Lalent factors and Measures	Data type	Loading	AME	Signal	Noise
Child health - Round 2, age 12					
1. Height for age z-score	Continuous	1	0.758	0.577	0.423
2. Weight	Continuous	1.258	0.954	0.925	0.075
3. How is child health?	Binary	0.886	0.111	–	–
Parental cognitive skills					
Mother's years of education	Continuous	1	0.818	0.674	0.326
Father's years of education	Continuous	0.986	0.806	0.658	0.342
Parental noncognitive skills					
1. Self-esteem score	Continuous	1	0.471	0.522	0.478
2. Self-efficacy score	Continuous	0.141	0.066	0.022	0.978
3. Self-respect and inclusion score	Continuous	0.757	0.356	0.266	0.734
Parental Investments					
1. Expenditure on the Young Lives child	Continuous	1	1.001	0.995	0.005
2. Number of hours studying outside school	Continuous	0.399	0.400	0.161	0.839
3. Quality of relationship	Continuous	0.067	0.067	0.009	0.991

6.2 Determinants of Parental Investments in Children

Parental investments reflect parental choices and decisions. The parental investment choices depend on their belief in child development, resources and preferences. The investment equation is a function of the child's cognitive and concognitive skills and health in the

[6] The estimated factor distribution moments, including factor means, standard deviations, correlations and mixture components are shown in Appendix C.

previous period, parental cognitive and noncognitive abilities and household characteristics.

Table 7 reports the estimates of the parental investment equation. Parental perception reflected by their cognitive and noncognitive abilities and their resources reflected by the wealth index have large effects on parental investment choices in their children. A 10% increase of a standard deviation in parental cognitive ability would increase parental investments in the child by 4.02% of a standard deviation, while a 10% increase of a standard deviation in parental noncognitive skills would increase investments by 1.72% of a standard deviation. This shows that better parental cognitive and noncognitive skills lead to better intergenerational skill outcomes through the investment channel. The wealth index has a positive, large and significant effect on investments; increasing the wealth index by a 10% would increase parental investments in children by 1.77 % of a standard deviation. Child health has a positive and significant effect on parental investments; a 10% increase of a standard deviation in child health would increase investments by 2.18% of a standard deviation.

Surprisingly, the child's cognitive and noncognitive skills do not have any impact on parental investments. It reflects that parents may not be aware of their child's skills and have too low expectations of returns to investments in skills. Attanasio et al. (2019) show that parents have distorted views about the child development process.

Parents in urban areas invest more in their children than parents in rural areas. Investments in female children are higher than investments in male counterparts. The number of siblings does not affect investments significantly.

The results also show that an increase in prices, except the price of notebook would increase investments, while shocks have negative effects on investments, except shock because of storms. In the sence of investments, the coefficients on prices should be negative. However, the goods considered are necessities in a relatively low-resource setting. Furthermore, I exploit the special (regional) variation in prices at the community level which is assumed to be not driven by demand differences and in a complex model with some alternative investment goods being complements to the goods used to proxy investment in the measurement system. Therefore, the estimated coefficients could be positive.

Table 7: Estimates of Parental Investment Function

Variables	Parental investments
Child's cognitive skills, age 12	-0.024
	(0.030)
Child's noncognitive skills, age 12	-0.102
	(0.076)
Child's health, age 12	0.218***
	(0.032)
Parental cognitive skills	0.402***
	(0.057)
Parental noncognitive skills	0.172***
	(0.066)
Wealth index	0.177***
	(0.045)
Female	0.080***
	(0.029)
Urban	0.530***
	(0.049)
Number of siblings aged 0-18	-0.032***
	(0.012)
Shock in input prices	-0.075**
	(0.034)
Drought	0.058**
	(0.029)
Flood	-0.060**
	(0.025)
Crops failure	-0.079***
	(0.022)
Illness of child's father	-0.006
	(0.047)
Illness of child's mother	-0.084**
	(0.041)
Illness of other household members	-0.114***
	(0.041)
Storm	0.118**
	(0.050)

Continued on the next page

Table 7 – *Continued from the previous page*

Variables	Parental investments
Price of notebook	0.038
	(0.024)
Price of clothes	0.250***
	(0.049)
Price of food	0.673***
	(0.136)
Price of medicine	0.069***
	(0.024)
Constant	-4.373***
	(0.654)
Observations	*961*

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.

6.3 Production Functions

Table 8 presents the production function estimates for cognitive skills, noncognitive skills and health. To characterize more cohesively the size and significance of the overall effects of each input in the production functions and assess the sensitivity of our major inputs, I analyze their marginal effects and recreate the dynamic process by using the estimated parameters.[7] First, I calculate the marginal effect to access the overall effects of each main input (Table 9). Second, I explore the role of self-productivity and cross-productivity in producing skills and health. Third, I consider the dynamic complementarity between skills, health and parental investments. My results show several essential features of human capital accumulation.

[7] The marginal effects are derived in Appendix D.

Table 8: Estimates of Production Functions

	Cognitive skills at age 15	Noncognitive skills at age 15	Health at age 15
	(1)	(2)	(3)
Child's cognitive skills at age 12	0.613***	0.001	0.064**
	(0.044)	(0.039)	(0.029)
Child's noncognitive skills at age 12	0.145***	0.339***	0.130**
	(0.055)	(0.077)	(0.053)
Child's health at age 12	0.042*	0.065	0.670***
	(0.025)	(0.053)	(0.073)
Parental Investments	0.261***	0.505***	0.346**
	(0.075)	(0.167)	(0.170)
Parental cognitive skills	0.017	-0.005	-0.184
	(0.041)	(0.122)	(0.193)
Parental noncognitive skills	-0.079	0.096	-0.026
	(0.050)	(0.081)	(0.065)
A_t	-0.010	0.211***	0.205***
	(0.024)	(0.031)	(0.024)
Control Function	-0.126*	-0.132	-0.420**
	(0.068)	(0.262)	(0.179)
Female	0.062***	-0.036	-0.238***
	(0.016)	(0.026)	(0.027)
Urban	-0.021	-0.328*	-0.421***
	(0.059)	(0.194)	(0.133)
Number of siblings aged 0-18	0.010	0.038***	0.007
	(0.009)	(0.011)	(0.009)
Complementarity(ρ)	-0.168	-1.851**	0.034
	(0.117)	(0.757)	(0.327)
Elasticity of substitution	0.856***	0.351*	1.036***
	(0.141)	(0.211)	(0.195)
Observations		*961*	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Marginal Effects

	Cognitive skills at age 15 (1)	Noncognitive skills at age 15 (2)	Health at age 15 (3)
Child's cognitive skills at age 12	0.614***	0.002	0.064**
	(0.043)	(0.039)	(0.029)
Child's noncognitive skills at age 12	0.145***	0.316***	0.130**
	(0.055)	(0.076)	(0.052)
Child's health at age 12	0.042*	0.085	0.670***
	(0.025)	(0.062)	(0.065)
Parental investments	0.261***	0.513***	0.346**
	(0.075)	(0.162)	(0.169)
Parental cognitive skills	0.017	-0.008	-0.184
	(0.042)	(0.123)	(0.194)
Parental noncognitive skills	-0.079	0.093	-0.026
	(0.050)	(0.078)	(0.067)
Observations		*961*	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.

6.3.1 Cognitive Skills

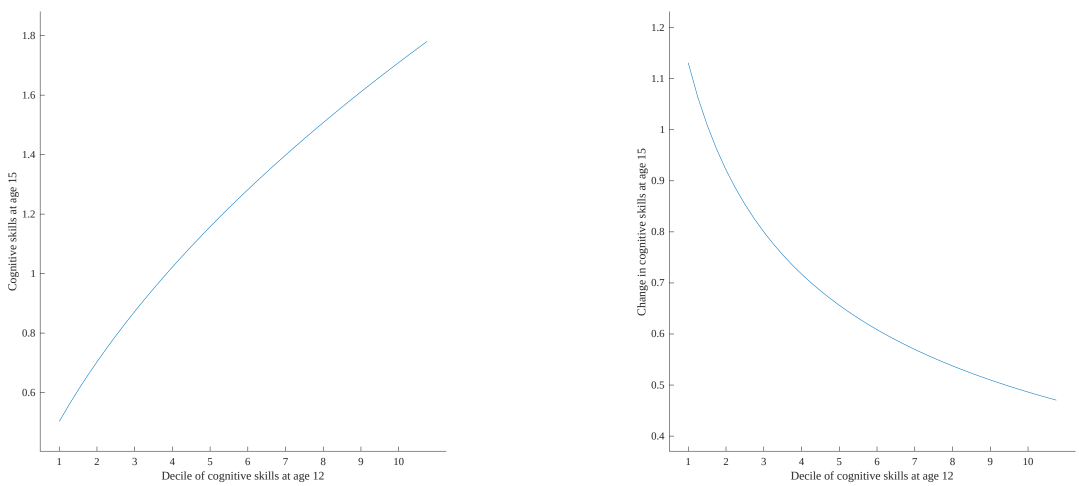
Table 8, Column 1 presents the estimates of the production function for cognitive skills and Table 9, Column 1 shows the marginal effects of main inputs. The coefficient on the investment residuals (control function) is significant and negative. It implies that investments are endogenous in the production of cognitive skills. The negative sign of this coefficient suggests that parents tend to increase their investments to compensate for an adverse shock that is unobserved but perceived by the parents and causes a decline in the child's cognitive skills. Ignoring this effect could lead to an underestimate of the impact of investments. Appendix E, Table E.1 and E.2 show that when investments are taken as exogenous, the coefficient on the investments is much lower, while the coefficients on other inputs are not dramatically affected. This result shows a compensatory role of parents to shocks to the child. The results show several other features of cognitive skill accumulation.

First, cognitive skills show a very strong self-productivity effect. That is, past cognitive skills have a strong and positive effect on current cognitive skills. Increasing cognitive skills at age 12 by 10% of a standard deviation would increase cognitive skills at age 15 by 6.13% of a standard deviation. Figure 1a illustrates the level of cognitive skills in the current period for each decile of past levels of cognitive skills, keeping all other inputs at their mean values. It shows that high cognitive skills produce high future cognitive skills. Figure 1b displays the self-productivity of cognition ($\partial\Theta_t^C/\partial\Theta_{t-1}^C$) for each decile of the levels of cognitive skills in the last period. These figures show that marginal increments of past cognitive skills are very productive ($\partial\Theta_t^C/\partial\Theta_{t-1}^C$ is high and positive for the entire space). It also demonstrates that the productivity of cognitive skills is higher for lower cognitive skill levels.

Figure 1: Cognitive Skills: Self-productivity

(a) Θ_t^C as a Function of Θ_{t-1}^C

(b) Cognitive Skills: Self-Productivity ($\partial\Theta_t^C/\partial\Theta_{t-1}^C$)



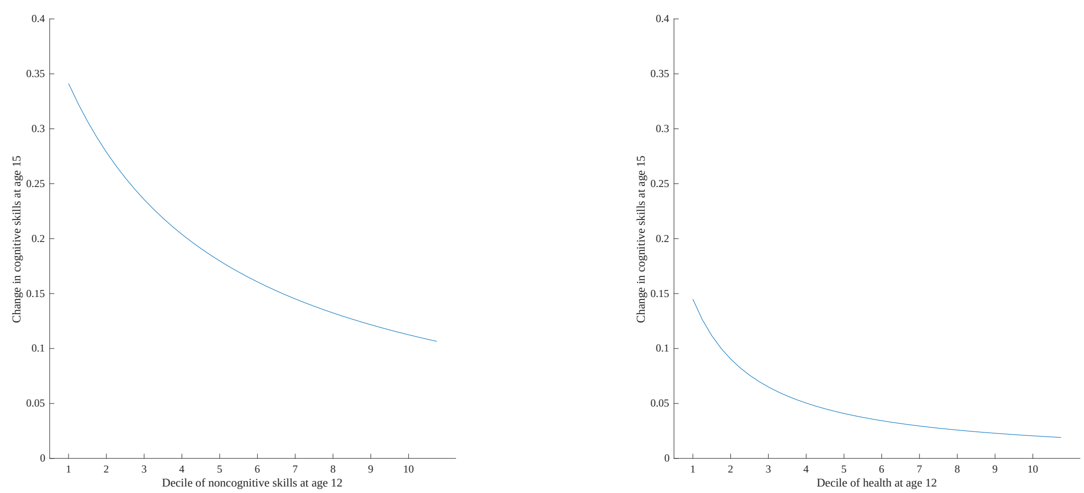
Second, there is cross-productivity from noncognitive skills and health to cognitive skills. That is, noncognitive skills and health in the last period foster the development of current cognitive skills. The cross-effect of noncognitive skills on cognition is strong and the cross-effect of health on cognition is small compared to that of noncognitive skills. A 10% increase of a standard deviation in the stocks of noncognitive skills and health in the last period would increase the current level of cognitive skills by 1.45% and 0.42% of a standard deviation respectively. Figure 2 shows the extent to which noncognitive skills and health affect the development of cognitive skills. $\Theta_t^C/\partial\Theta_{t-1}^{NC}$ and $\Theta_t^C/\partial\Theta_{t-1}^{H}$ are positive for the entire space and the cross-effect of noncognitive skills is stronger than that of health. The largest impacts are for children with the lowest level of noncognitive skills and health. These results demonstrate the importance of noncognitive skills and good health in developing cognitive skills.

These results are consistent with previous studies (Cunha et al., 2010; Helmers and Patnam, 2011; Attanasio et al., 2017; Sánchez, 2017). Cunha et al. (2010) and Helmers and Patnam (2011) find that self-productivity of cognition and cross-productivity from noncognitive skills to cognitive skills play an important role in the formation of skills with larger level effects of self-productivity of cognition. The result of cross-productivity effects of health on cognition is aligned with that of Attanasio et al. (2017) and Sánchez (2017) which indicates that health is important for future cognitive skill development.

Figure 2: Cognitive skills: Cross-productivity

(a) Cross-productivity from Noncognitive Skills to Cognitive Skills ($\partial\Theta_t^C/\partial\Theta_{t-1}^{NC}$)

(b) Cross-productivity from Health to Cognitive Skills ($\partial\Theta_t^C/\partial\Theta_{t-1}^{H}$)



Figures 3, 4 and 5 show that the self-effect of cognitive skills and cross-effects of noncognitive skills and health in producing future cognitive skills are higher for those who have higher levels of cognitive skills, noncognitive skills and health. That is, at the same initial stock of cognition, the self-effect is higher for those with higher initial noncognitive skills and better health. Similarly, with the same initial level of noncognitive skills, the cross-effect from noncognitive skills to cognitive skills is higher for those with higher initial cognition and better health, and the cross-effect from health to cognitive skills is higher for those with higher initial stocks of cognitive and noncognitive skills. These results mean that higher stocks of a certain dimension of human capital make the marginal increments of other dimensions of human capital more productive.

Figure 3: Cognitive Skills: Self-productivity ($\partial\Theta_t^C/\partial\Theta_{t-1}^C$)

(a) (b)

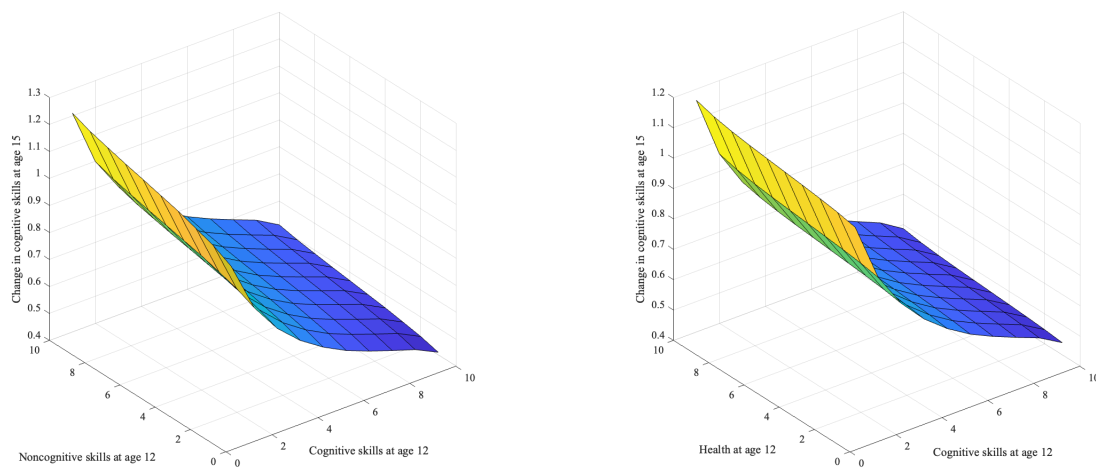
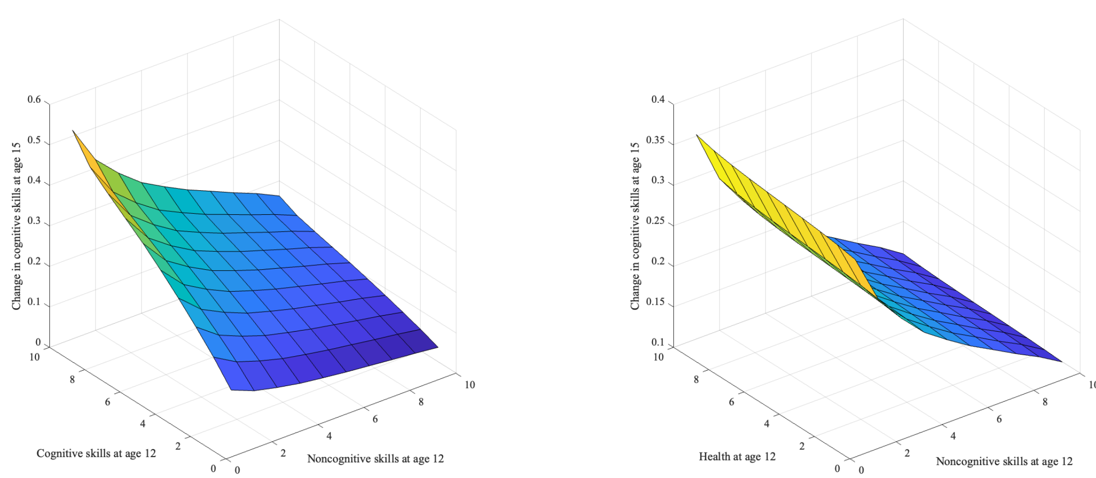


Figure 4: Cognitive Skills: Cross-productivity from Noncognitive Skills to Cognitive Skills ($\partial\Theta_t^C/\partial\Theta_{t-1}^{NC}$)

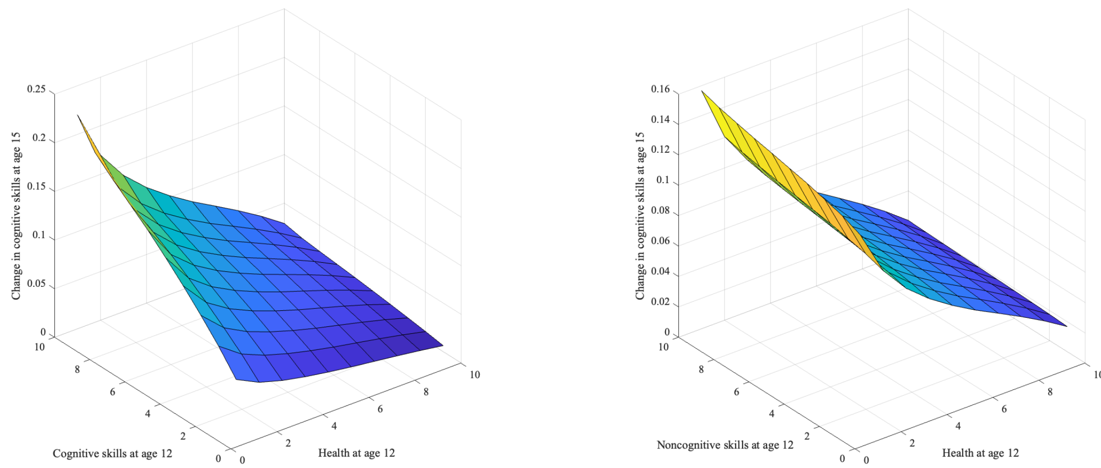
(a) (b)



Third, parental cognition and con-cognition do not have significant impacts on child cognitive skills. It indicates that parental background of cognitive and noncognitive skills only plays an important role in their investment decisions in children and indirectly develops their

Figure 5: Cognitive Skills: Cross-productivity from Health to Cognitive Skills ($\partial\Theta_t^C/\partial\Theta_{t-1}^H$)

(a) (b)



child's cognitive skills through investments. Girls are more likely to have higher cognitive skills than boys. The residence of the child and the number of siblings do not seem to have a significant effect on cognitive skill accumulation, although they matter for parental investment decisions.

Fourth, one of the key estimates is the complementarity coefficient ($\rho = -0.168$). The elasticities of substitution between the various inputs ($\sigma = 1/(1\text{-}\rho) = 0.856$) show some degree of complementarities. This result is aligned with the existing literature (Attanasio et al., 2020; Cunha et al., 2010). I also reject the hypothesis that $\rho = 1$, which indicates that the production function is linear and the inputs are additively separable. These results imply that it is not easy to compensate and remediate low levels or deficits of skills and health from the previous periods.

Finally, the key result, which largely motivates this study, is the role of parental investments. First, parental investments have a very strong effect on the child's cognitive skills. A 10% increase of a standard deviation in parental investments would increase cognitive skills by 2.61% of a standard deviation respectively. Second, to understand the extent to which parental investment can affect the accumulation of cognitive skills of the child, I explore the dynamic complementarity between cognitive skills and investments, a concept introduced in Cunha and Heckman (2007) to imply that past cognitive skills and past investments in those skills increase the productivity of current investments ($\partial^2\Theta_t^C/\partial\Theta_t^I\partial\Theta_{t-1}^C > 0$). Figure 6 shows a strong dynamic complementarity between cognitive skills and investments, meaning that returns to investments are higher for children with better initial cognitive skills or higher past cognitive skills make investments more productive. Figure 7 shows compounded complementarity effects under the effects of noncognitive skills and health, higher noncognitive skills and health make investments even more productive. These results is in line with the existing literature (Attanasio et al., 2020; Attanasio et al., 2017; Cunha and Heckman, 2007; Cunha et al., 2010).

Figure 6: Complementarity between Investments and Cognitive Skills

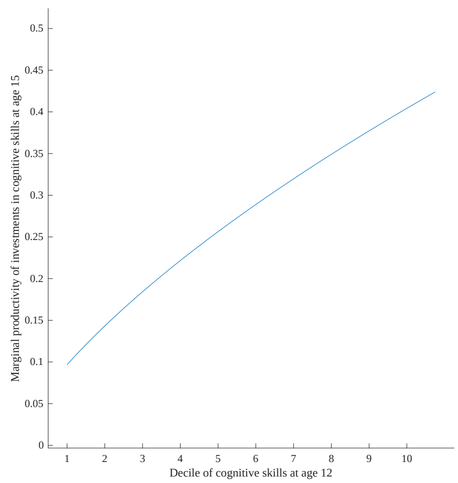
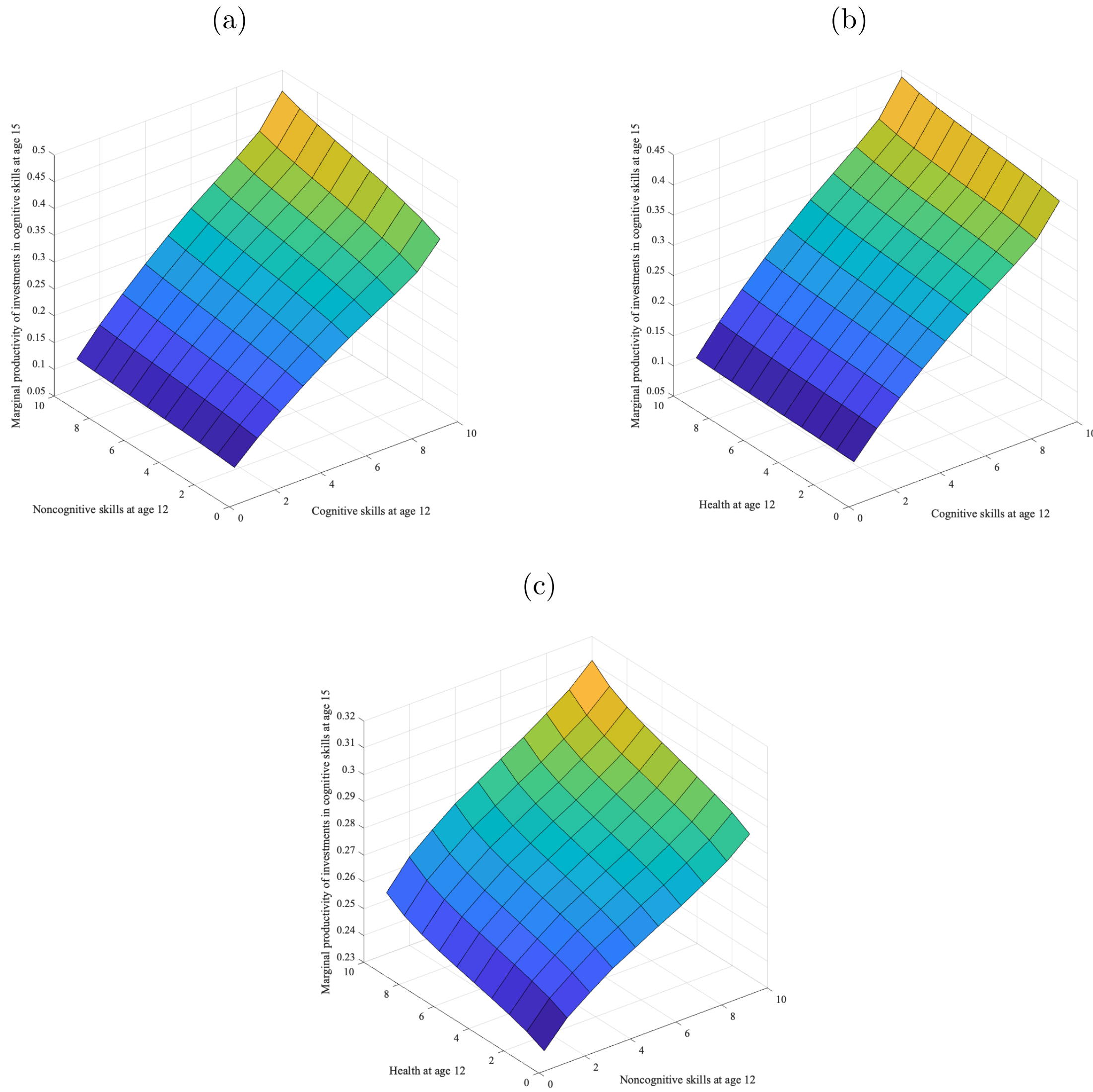


Figure 7: Complementarity between Investments and Cognitive Skills



The self-effects, cross-effects and dynamic complementarity together become a dynamic multiplier effect mechanism of cognitive skill accumulation whereby skills, health and investments produce cognitive skills. These multiplier effects would lead to different rates of cognitive skill growth for children with different initial skills and health. Furthermore, parental skills and resources are positively associated with parental investments, which implies that children with better family backgrounds get more investments and they use these investments more effectively. These effects could lead to substantial increases in inequality in producing skills and finally lead to social inequality. Dynamic complementarity is crucial since it could be a source of inequality and it shows the role of investments. These results also indicate the importance of interventions by boosting investments at early ages that can alter child development path, especially for disadvantaged children. A lack of parental investments can seriously hinder the development of a child.

6.3.2 Noncognitive Skills

The analysis of noncognitive skills follows the same methodology as the one for cognitive skills, presented in the previous section. I find no evidence of endogenous investments in noncognitive skill production. The results are shown in Table 8, Column 2 and Table 9, Column 2 confirming the evidence of self-productivity of noncognitive skills in which noncognitive skills in an earlier period produce better noncognitive skills in the next period. A 10% increase of a standard deviation in noncognitive skills at age 12 would increase noncognitive skills at age 15 by 3.16% of a standard deviation. Figure 8 shows that noncognitive skills are productive in inducing better noncognitive outcomes ($\partial\Theta_t^{NC}/\partial\Theta_{t-1}^{NC} > 0$) and the self-effects of noncognitive skills are higher for children with higher initial levels of cognition and health (Figure 9).

Figure 8: Noncognitive skills: Self-productivity

(a) Θ_t^{NC} as a Function of Θ_{t-1}^{NC}

(b) Noncognitive Skills: Self-productivity ($\partial\Theta_t^{NC}/\partial\Theta_{t-1}^{NC}$)

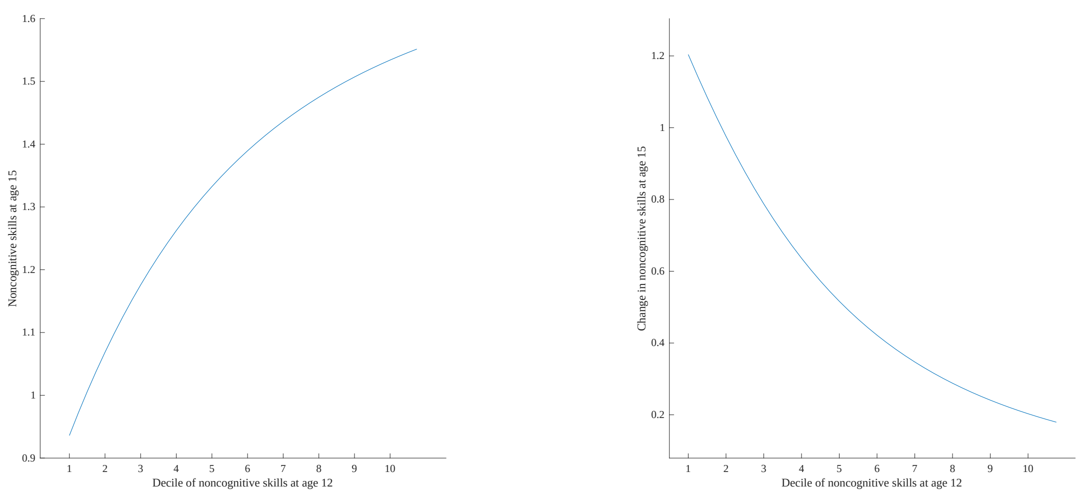
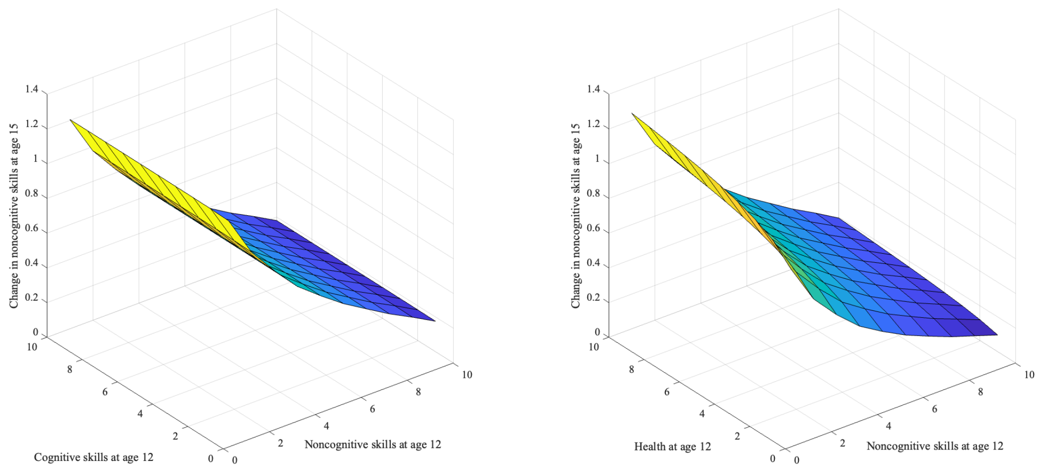


Figure 9: Noncognitive Skills: Self-productivity ($\partial \Theta_t^{NC} / \partial \Theta_{t-1}^{NC}$)

(a) (b)



The results do not show the existence of cross-productivity effects of cognitive skills and health on noncognitive skills. The initial levels of cognitive skills and health do not contribute to producing noncognitive skills. The cross-effects are similar among children with different initial levels of noncognitive skills except for the lowest deciles of cognition and health (Figures 10, 11 and 12).

Figure 10: Noncognitive Skills: Cross-productivity from Cognitive Skills and Health

(a) Cross-productivity from Cognitive Skills to Noncognitive skills ($\partial \Theta_t^{NC} / \partial \Theta_{t-1}^{C}$)

(b) Cross-productivity from Health to Noncognitive Skills ($\partial \Theta_t^{NC} / \partial \Theta_{t-1}^{H}$)

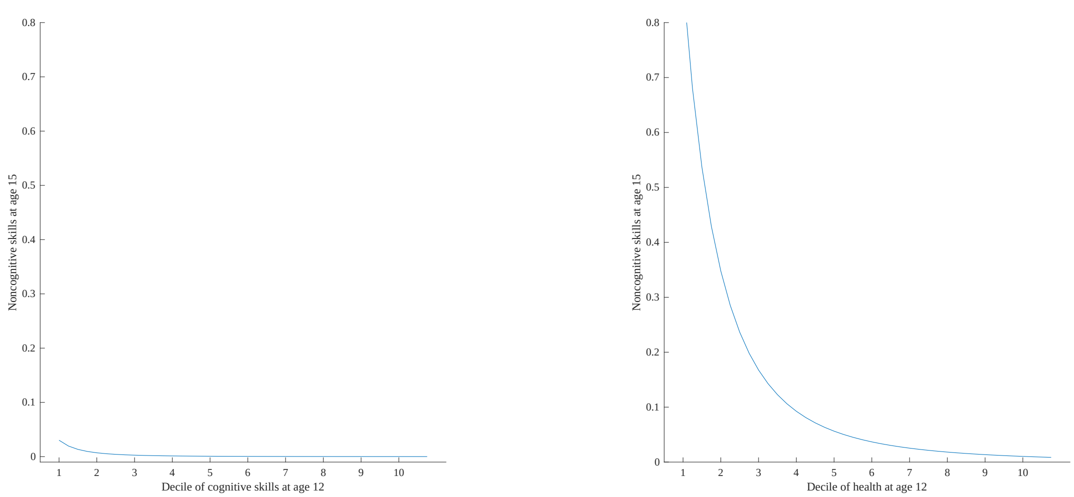


Figure 11: Noncognitive Skills: Cross-productivity from Cognitive Skills to Noncognitive Skills ($\partial \Theta_t^{NC} / \partial \Theta_{t-1}^{C}$)

(a) (b)

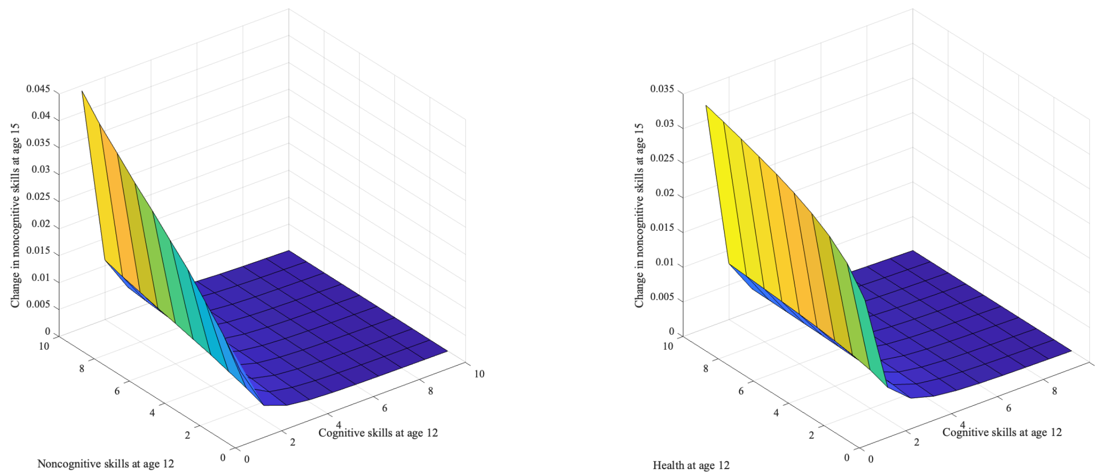
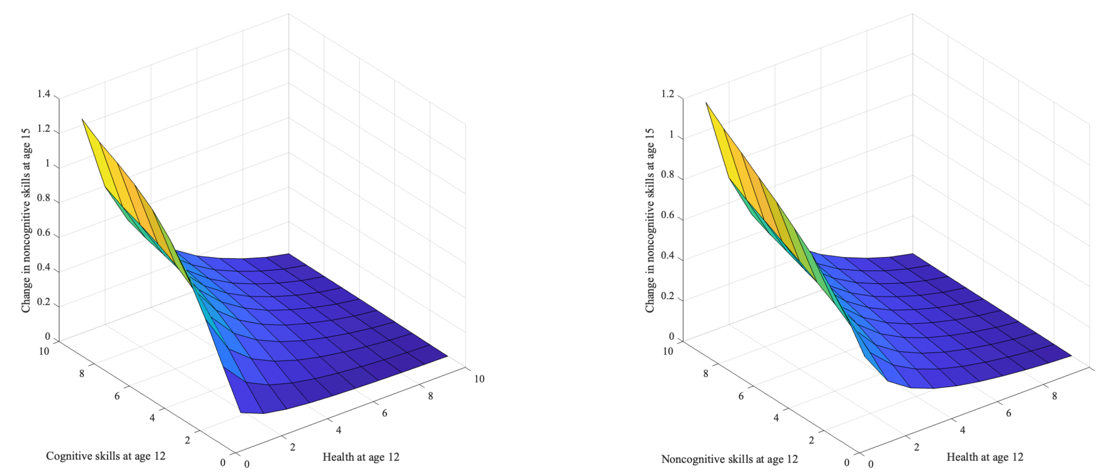


Figure 12: Noncognitive Skills: Cross-productivity from Health to Noncognitive Skills ($\partial \Theta_t^{NC} / \partial \Theta_{t-1}^{H}$)

(a) (b)

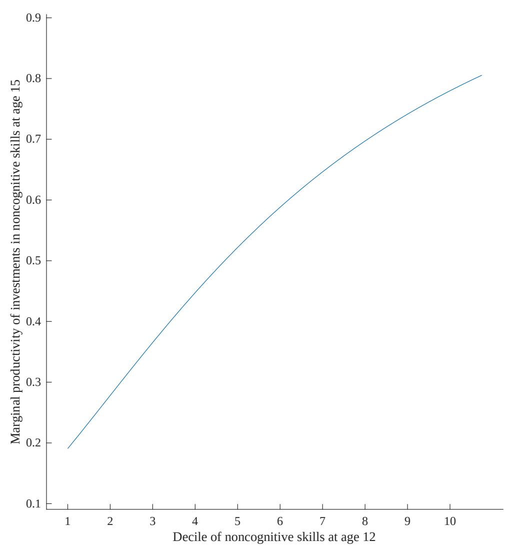


There is no evidence that parental cognitive and noncognitive skills have a direct effect on the child's noncognitive skills.

The complementarity coefficient (ρ = -1.851) and the elasticities of substitution (σ = 1/(1-ρ) = 0.351) show is relatively small compared to those of cognitive skills. This suggests the complementarity among inputs is stronger for the noncognitive production function and it is more difficult to remedy deficits of noncognitive skills by investments.

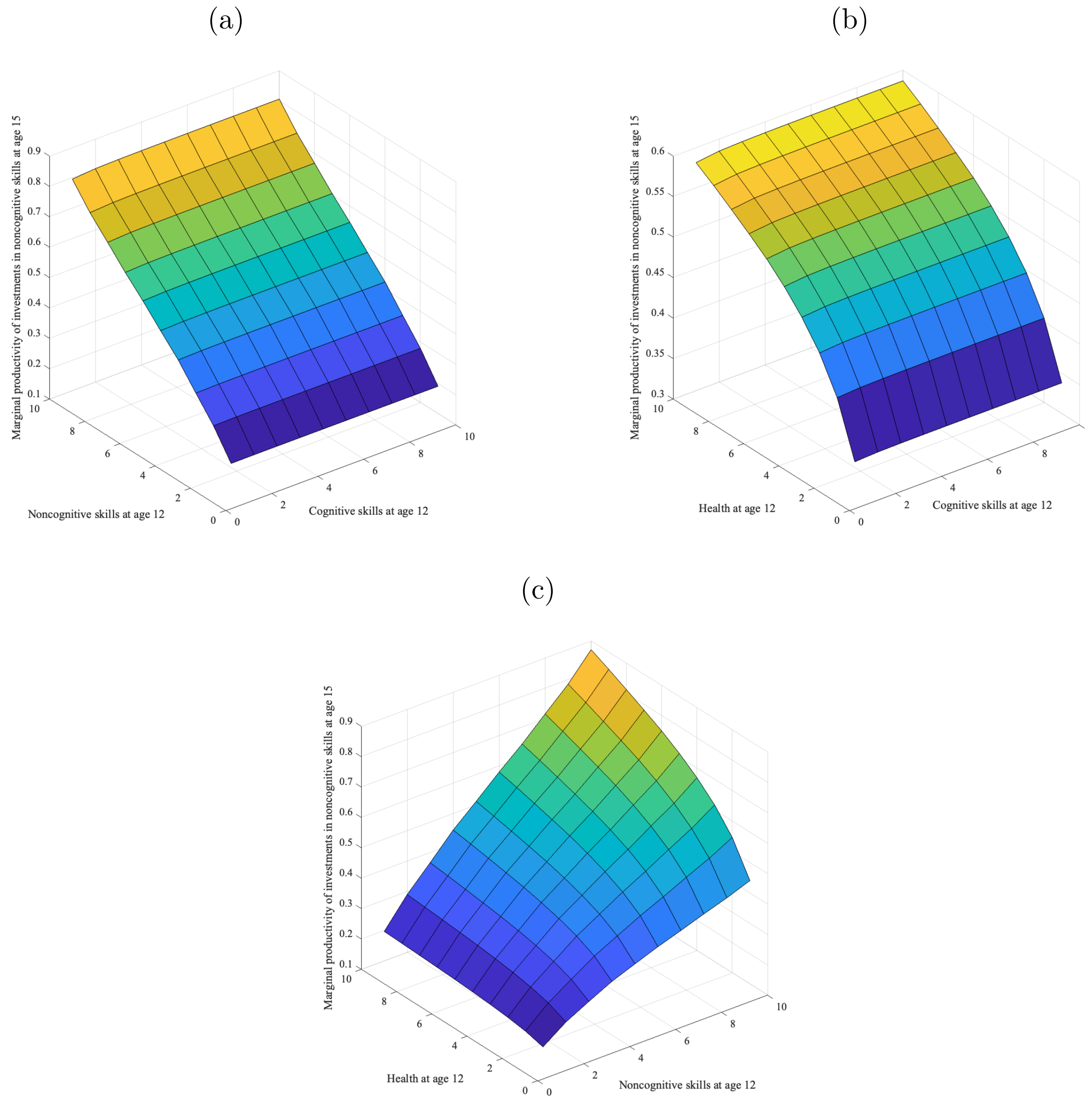
Parental investments have a very strong effect on a child's noncognitive skills. A 10% increase of a standard deviation in parental investments would increase noncognitive skills by 5.13% of a standard deviation. The effect of investments on noncognitive skills is larger than that on cognitive skills and health, which will be discussed in Section 6.3.3. Figures 13 and 14 show a strong dynamic complementarity between noncognitive skills and investments and the marginal effects of investments are higher among children with higher initial noncognitive skills. It is noted that child cognitive skills and health in the previous period do not have significant impacts on the current stock of noncognitive skills as discussed above.

Figure 13: Complementarity between Investments and Noncognitive Skills



My results share certain commonalities with the study of Cunha et al. (2010) for the US in the sense that noncognitive skills are self-productive, cognition is not cross-productive for noncognitive skill formation and investments are an important factor for noncognitive skill accumulation. However, the result about cross-productivity effects contrasts with that of Attanasio et al. (2020) and Sánchez (2017), which find that current cognitive skills foster future noncognitive skill accumulation. My result about self-productivity is consistent with that of Attanasio et al. (2020), Attanasio et al. (2017), Cunha et al. (2010)) and Sánchez (2017).

Figure 14: Complementarity between Investments and Noncognitive Skills



6.3.3 Child Health

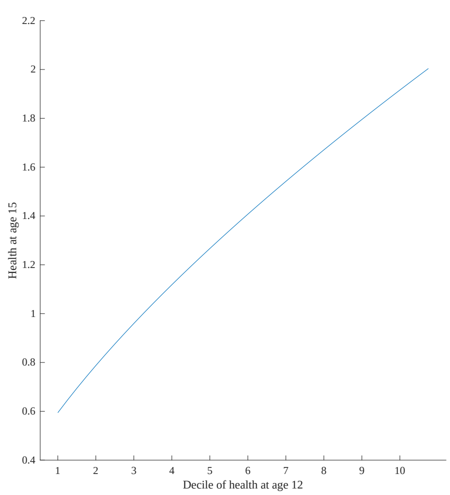
Now turn to the production of health. Table 8, Column 3 and Table 9, Column 3 present the estimates of the production process of health and the marginal effects of main inputs. As with the cognitive skill production function, the coefficient on the investment residuals is significant and negative. This implies that investments are endogenous in the production function of health and the negative sign suggests that parents tend to compensate for adverse shocks to their children by increasing investments.

The estimated results show strong evidence of self-productivity and the existence of cross-productivity from cognitive and noncognitive skills to health in the production process of health. A 10% increase of a standard deviation in health in the previous stage would increase health in the current stage by 6.7% of a standard deviation. Increasing child cognitive and noncognitive skills in the previous period by 10% of a standard deviation would increase health in the current period by 0.64% and 1.3% of a standard deviation respectively.

Figures 15 and 16 show the self-productivity of health and Figures 17, 18 and 19 show cross-productivity from cognitive skills and noncognitive skills to health. Figures 18 and 19 show the cross-effects of cognitive and noncognitive skills on health are higher for those with high initial levels of health and this productivity does not change much among deciles of cognitive and noncognitive skills.

Figure 15: Health: Self-productivity

(a) Θ_t^H as a Function of Θ_{t-1}^H



(b) Health: Self-Productivity ($\partial\Theta_t^H/\partial\Theta_{t-1}^H$)

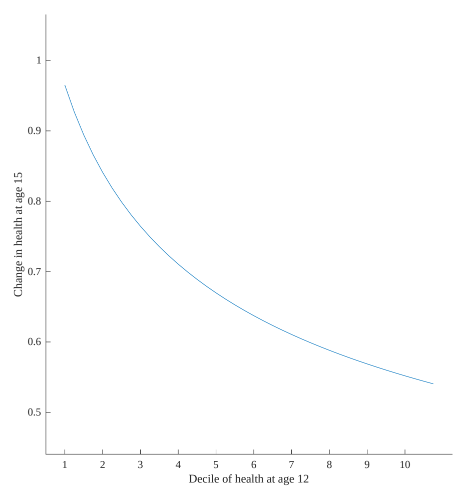


Figure 16: Health: Self-productivity ($\partial\Theta_t^H/\partial\Theta_{t-1}^H$)

(a) (b)

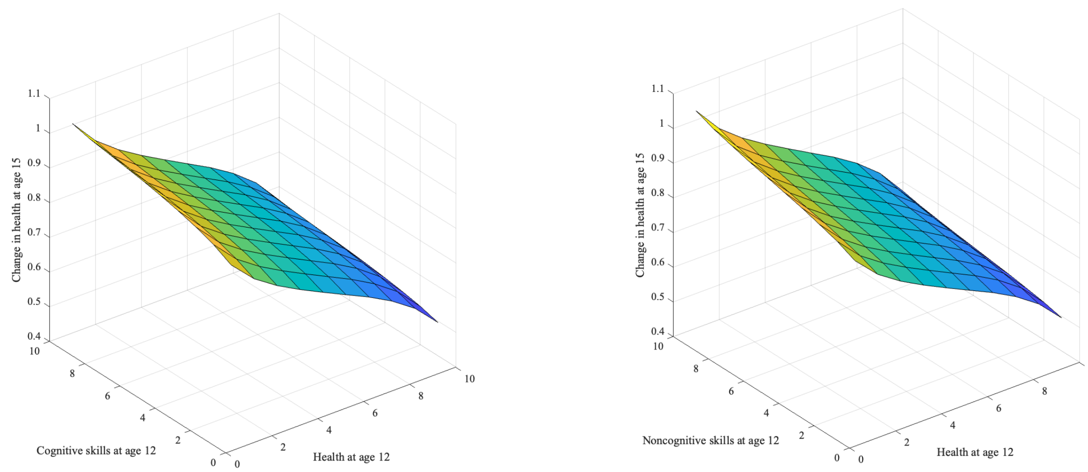


Figure 17: Health: Cross-productivity

(a) Cross-productivity from Cognitive Skills to Health($\partial\Theta_t^H/\partial\Theta_{t-1}^C$)

(b) Cross-productivity from noncognitive skills to health ($\partial\Theta_t^H/\partial\Theta_{t-1}^{NC}$)

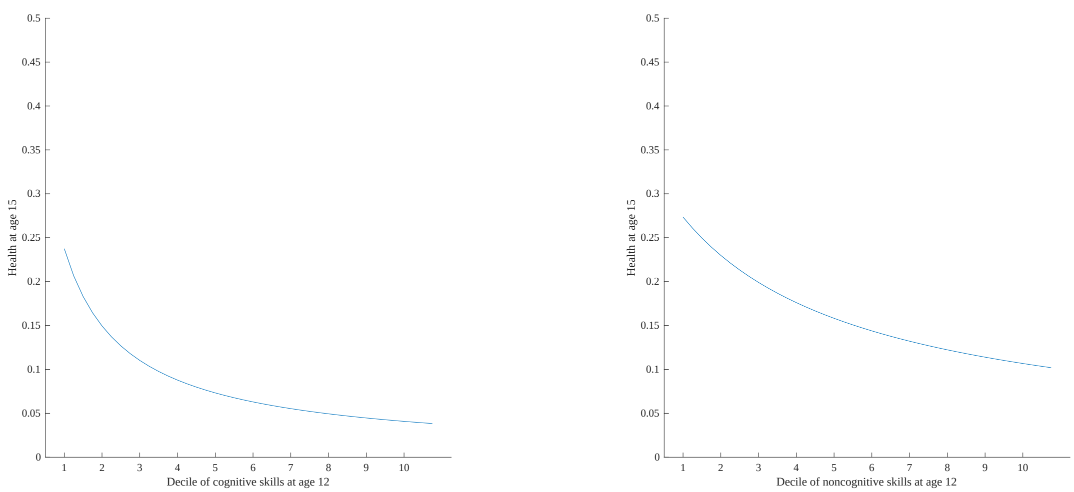


Figure 18: Health: Cross-productivity from Cognitive Skills to Health ($\partial \Theta_t^H / \partial \Theta_{t-1}^C$)

(a) (b)

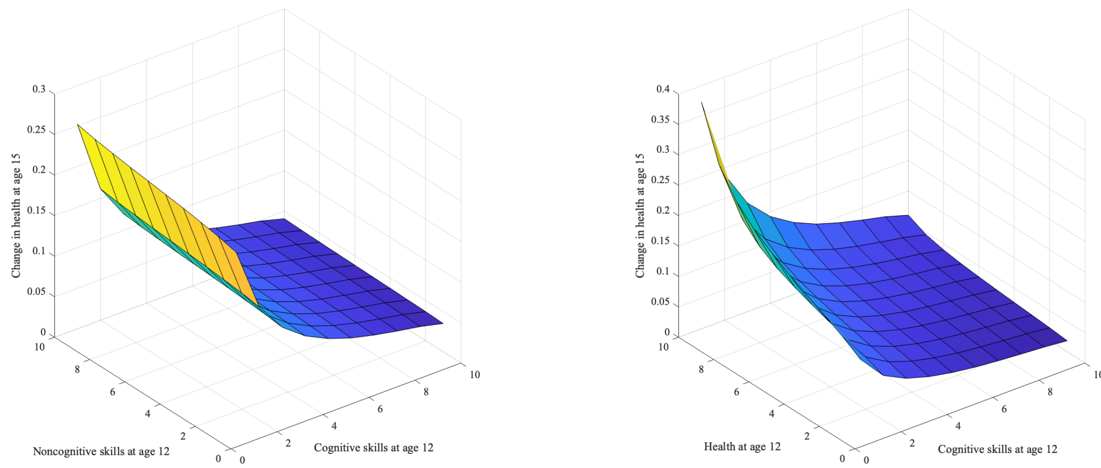
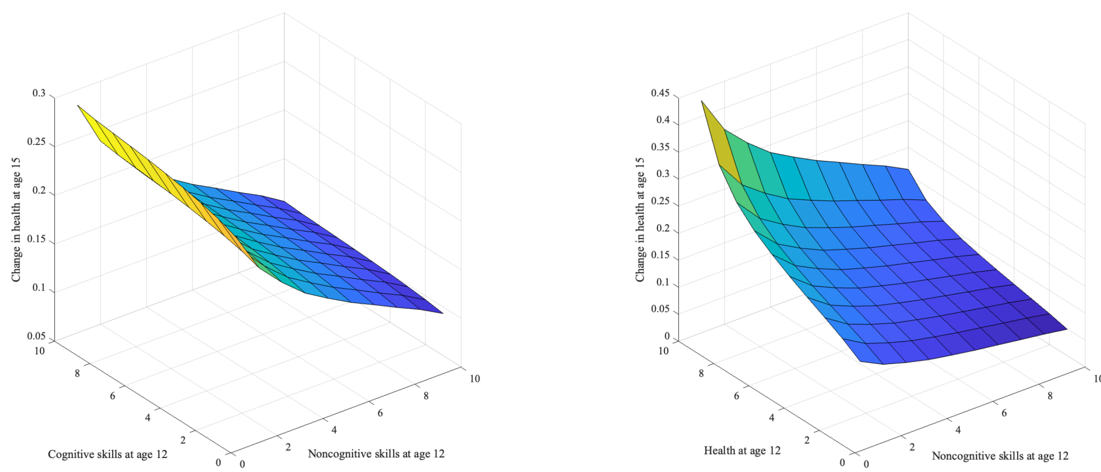


Figure 19: Health: Cross-productivity from Noncognitive Skills to Health ($\partial \Theta_t^H / \partial \Theta_{t-1}^{NC}$)

(a) (b)



The results also show a strong impact of investments on health and strong evidence of a dynamic complementarity between parental investments and their child's health. Table 9, Column 3 shows that A 10% increase of a standard deviation in parental investments would increase health by 3.46% of a standard deviation. . Figures 20 and 21 show strong complementarity between investments and child health. With the important roles of health on child development as discussed here and given that parental investments strongly affect child health, interventions that address health deficits as early as possible are critical for child development, especially in low-resource settings where health deficits at early ages are common.

While Attanasio et al. (2017) reported mixed results about self-productivity effects of cognitive skills on health, my result is aligned with them in terms of self-productivity effects and large impacts of investments.

Figure 20: Complementarity between Investments and Child Health

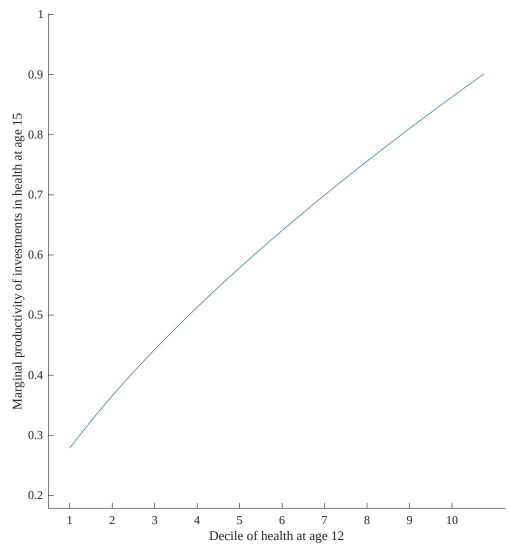
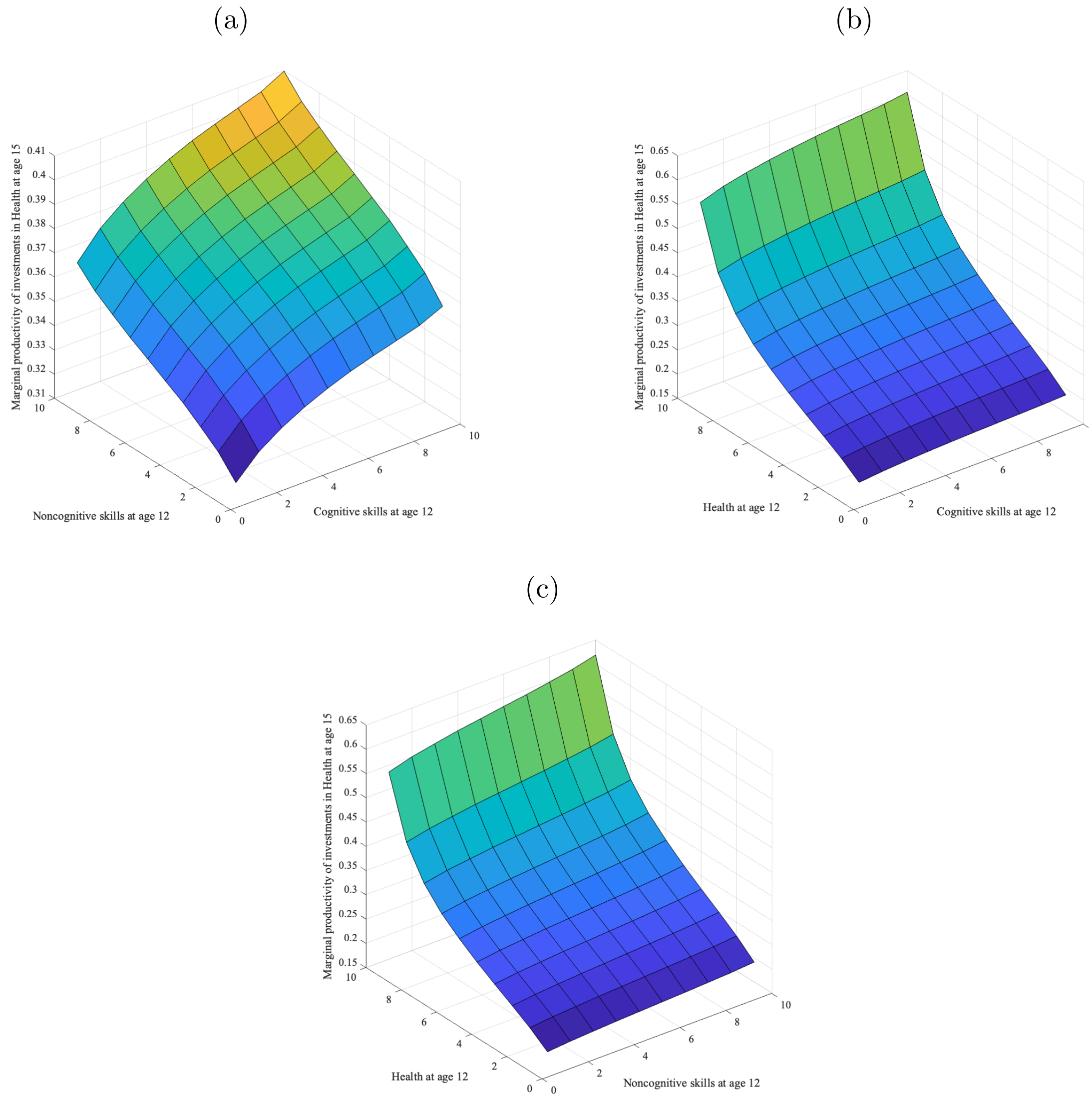


Figure 21: Complementarity between Investments and Child Health



Parental cognitive and noncognitive skills have no impact on health. Being a boy, living in rural area and having fewer siblings have positive impacts on health.

7 Conclusion

Understanding the evolution of human capital, how its various constituents (cognitive skills, noncognitive skills and health) interact in childhood and its intergenerational transmission is vital to design policies and interventions that can improve productivity and contribute to sustainable economic growth and development. This is especially true in the context of developing countries where children face various developmental risk factors and health deficits that deter their development. With rich and unique data from the Young lives survey that contains multiple dimensions and indicators of cognitive skills, noncognitive skills and health, this study uses the dynamic factor analysis approach to recover latent factors that can more precisely capture the multidimensional nature of abilities by lower dimensional factors and correct for measurement errors. This study estimates the endogenous parental investment function jointly with the dynamic model of cognitive skills, noncognitive skills and health.

The results show that self-effects are present and strong in the production of all human capital dimensions. That is, skills produce skills and health produces health. The results also confirm the existence of cross-effects, except cross-productivity from cognitive and health to noncognitive skills: the existing stock of noncognitive skills and health foster the production process of cognitive skills; high cognitive and noncognitive skills lead to better health; and cognitive skills and health are unimportant to noncognitive skill development. These results indicate that there is a high cost for the accumulation of human capital for those who start with lower skill and health levels.

Most importantly, the results confirm the vital importance of parental investments. First, investments strongly and directly affect the accumulation of skills and health. Their impacts are the largest for noncognitive skills, followed by health and cognition. Second, there is a dynamic complementarity among the inputs in human capital production. This implies that returns to investments are higher for children with better initial conditions. Furthermore, it also implies that higher initial stocks of skills and health make those skills and health more productive.

The self-productivity, cross-productivity and dynamic complementarity together become a dynamic multiplier effect mechanism of skill and health accumulation whereby skills, health and parental investments produce skills and health. Furthermore, parental investment decisions strongly depend on parental skills and wealth. This indicates that children with better backgrounds get more investments, and they can also use these investments more productively. These effects could lead to substantially different growth rates of human capital and substantial increases in inequality in producing skills and finally lead to social inequality.

The results provide some insight as to how parents make investment choices in their children and show the role of parental investments as a source of child development and inequality. The findings also indicate the importance of interventions by boosting investments at early ages that can alter child development path, especially for disadvantaged children. A lack of parental investments can seriously hinder the development of a child.

My results provide important evidence that skills and health are produced from a combination of an individual's skills and health, parental skills and investments and other individual and family factors. Therefore, policies and interventions to develop human capital need to take into account of the complex interactions over childhood among these factors. Policies, interventions and investments in children at early ages are key to improving skills and health deficits in human development and contributing to social inequality reduction.

References

Agostinelli, F. and M. Wiswall (2016). Estimating the technology of children's skill formation. Technical report, National Bureau of Economic Research.

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review 110*(1), 48–85.

Attanasio, O., F. Cunha, and P. Jervis (2019). Subjective parental beliefs. their measurement and role. Technical report, National Bureau of Economic Research.

Attanasio, O., C. Meghir, E. Nix, and F. Salvati (2017). Human capital growth and poverty: Evidence from Ethiopia and Peru. *Review of Economic Dynamics 25*, 234–259. Special Issue on Human Capital and Inequality.

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association 13*(6), 949–997.

Aucejo, E. and J. James (2021). The path to college education: The role of math and verbal skills. *Journal of Political Economy 129*(10), 2905–2946.

Coneus, K., M. Laucht, and K. Reuß (2012). The role of parental investments for cognitive and noncognitive skill formation—evidence for the first 11 years of life. *Economics and Human Biology 10*(2), 189–209.

Crookston, B., R. Forste, C. McClellan, A. Georgiadis, and T. Heaton (2014). Factors associated with cognitive achievement in late childhood and adolescence: The young lives cohort study of children in ethiopia, india, peru, and vietnam. *BMC pediatrics 14*, 253.

Cunha, F. and J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97*(2), 31–47.

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

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education 1*, 697–812.

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

Del Boca, D., C. Flinn, and M. Wiswall (2013). Household choices and child development. *The Review of Economic Studies 81*(1), 137–185.

Dercon, S. and P. Krishnan (2009). Poverty and the psychosocial competencies of children: evidence from the young lives sample in four developing countries. *Children Youth and Environments 19*(2), 138–163.

Doyle, O., C. P. Harmon, J. J. Heckman, and R. E. Tremblay (2009). Investing in early human development: Timing and economic efficiency. *Economics and Human Biology 7*(1), 1–6.

Engle, P. L., M. M. Black, J. R. Behrman, M. C. De Mello, P. J. Gertler, L. Kapiriri, R. Martorell, M. E. Young, I. C. D. S. Group, et al. (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The lancet 369*(9557), 229–242.

Glewwe, P., Q. Huang, and A. Park (2017). Cognitive skills, noncognitive skills, and school-to-work transitions in rural China. *Journal of Economic Behavior & Organization 134*, 141 – 164.

Heckman, J. J. (2008). The case for investing in disadvantaged young children. *CESifo DICE Report 6*(2), 3–8.

Heckman, J. J., J. Stixrud, and S. Urzúa (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

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

James, J. (2017). MM algorithm for general mixed multinomial logit models. *Journal of Applied Econometrics 32*(4), 841–857.

Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of cognitive neuroscience 16*(8), 1412–1425.

Knudsen, E. I., J. J. Heckman, J. L. Cameron, and J. P. Shonkoff (2006). Economic, neurobiological, and behavioral perspectives on building america's future workforce. *Proceedings of the national Academy of Sciences 103*(27), 10155–10162.

Nguyen, N. (2008, 01). An assessment of the young lives sampling approach in vietnam. *University of Oxford, Open Access publications from University of Oxford*.

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *Journal of economic Perspectives 4*(4), 25–45.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton University Press.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied 80*(1), 1.

Sánchez, A. (2017). The structural relationship between early nutrition, cognitive skills and non-cognitive skills in four developing countries. *Economics and Human Biology 27*, 33–54.

Urzúa, S. (2008). Racial labor marketgaps: The role of abilities and schooling choices. *The Journal of Human Resources 43*(4), 919–971.

Walker, S. P., S. M. Chang, C. A. Powell, and S. M. Grantham-McGregor (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted jamaican children: prospective cohort study. *The lancet 366*(9499), 1804–1807.

Appendix A: The Construction of Measures of Child's and Parental Noncognitive Skills and Quality of Relationship

Table A.1: Construction of Measures of Noncognitive Skills and Quality of Relationship

Scale/Index	**Items/Statements**
Child self-esteem - Round 3, age 15	The self-esteem scale is constructed using the following items/Statements: 1. 'I am proud of my clothes'; 2. 'I feel my clothing is right for all occasions'; 3. 'I am proud of my shoes or of having shoes'; 4. 'I am proud because I have the right books, pencils or other equipment for school'; 5. 'I am proud that I have the correct uniform'; 6. 'I am proud of the work I have to do'.
Child self-esteem - Round 2, age 12	The items are: 1. 'I am proud of my clothes'; 2. 'I am proud of my shoes or of having shoes'; 3. 'I am proud because I have the right supplies for school'; 4. 'I am proud that I have the correct uniform'; 5. 'I feel proud to show my friends where I live'. 6. 'I feel proud of the job done by the head of household'. 7. 'I am proud of my achievements at school'.

Continued on the next page

Table A.1: The Construction of Measures *Continued*

Scale	Items/Statements
Child self-efficacy - Round 3, age 15	The items are: 1. 'If we try hard we can improve my situation in life'; 2. 'Other people in my family make all the decisions about how we spend my time';* 3. 'I like to make plans for my future studies and work'; 4. 'If we study hard we will be rewarded with a better job in the future'; 5. 'I have a choice about the work I do'.
Child self-efficacy - Round 2, age 12	The items are: 1. 'If we try hard we can improve my situation in life'; 2. 'Other people in my family make all the decisions about how we spend my time';* 3. 'I like to make plans for my future studies and work'; 4. 'If we study hard we will be rewarded with a better job in the future';
Child self-respect and Inclusion - Round 3, age 15	1. 'When I am at the shops/market I am usually treated by others with fairness and respect'; 2. 'Adults in my community treat me as well as they treat other children of my age'; 3. 'The other children in my class treat me with respect'; 4. 'Other pupils in my class tease me at school'; 5. 'My friends will stand by me during difficult times'; 6. 'I feel I belong at my school'; 7. 'My friends look up to me as a leader'; 8. 'I have people I look up to' 9. 'I have opportunities to develop job skills'.

Continued on the next page

Table A.1: The Construction of Measures *Continued*

Scale	**Items/Statements**
Child self-respect and Inclusion - Round 2, age 12	1. ‘At the shops I am treated with fairness’; 2. ‘Adults in my street treat me worse than other children of my age’[*]; 3. ‘The other children in my class treat me with respect’; 4. ‘Other pupils in my class tease me at school’[a]; 5. ‘My teachers treat me worse than other children’[*];
Parental self-esteem - measured at Round 2	1. ‘ I feel proud to show my friends or other visitors where I live’; 2. ‘I am ashamed of my clothes’[*]; 3. ‘I feel proud of the job done by the household head’; 4. ‘The job I do makes me feel proud’; 5. ‘I feel proud of my children;
Parental self-efficacy - measured at Round 2	The items are: 1. ‘If we try hard we can improve my situation in life’; 2. ‘I like to make plans for my future studies and work’; 3. ‘I have no choice about which school to send my child to’[*]; 4. ’If my child gets sick I can do little to help him/her get better’[*]; 5. ’I can do little to help my child do well in school no matter how hard I try’[*];
Parental self-respect and Inclusion - measured at Round 2	1. ‘At the shops I am treated with fairness and respect’; 2. ‘Other people in the street look down on me and my family’[*]; 3. ‘My children’s teachers are unfriendly or rude to me’[*];

Continued on the next page

Table A.1: The Construction of Measures *Continued*

Scale	Items/Statements
Quality of relationship - Round 3	1. 'I always obey my parents' ; 2. 'My parents rarely talk to me about the things that matter to me';* 3. 'I always feel loved by my parents'; 4. ' My parents never support me in the things I want to do';* 5. 'I usually feel able to speak my views and feelings with my parents'; 6. Most of the time my parents treat me fairly when I do something wrong'; 7. 'Compared to my sisters fewer things are provided for me';* 8. 'I receive lots of time and attention from my parents'; 9. 'Compared to my brothers fewer things are provided for me';* 10. 'Compared to my brothers I have less freedom to leave the house when I want';* 11. 'Compared to my sisters I have less freedom to leave the house when I want';* 12. 'My parents treat me worse than other children in my family';*

* The item is recoded to reflect a positive statement

Appendix B: Model Estimation Procedure

The full log-likelihood function for the measurement system I want to estimate is Equation 12:

$$\mathcal{L}(\Psi) = \sum_{i=1}^{N} ln\left(\int L(T_i|\theta_i)f(\theta)d\theta\right) \tag{B.1}$$

Where Ψ are all the parameters of the model that I want to estimate, $\Psi = \{\alpha, \beta, \sigma, \tau_c, \mu_c, \Omega_c\}$.

I maximize the log-likelihood function B.1 using the minorization-maximization algorithm developed in James (2017).

Given an initial value of parameters, Ψ^0, the log-likelihood function $\mathcal{L}(\Psi)$ can be bounded below by a quadratic function:

$$Q(\Psi|\Psi^0) = \sum_{i=1}^{n} \int ln(L(T_i|\theta_i)f(\theta))h(\theta|T_i, \Psi^0)d\theta \tag{B.2}$$

where

$$h(\theta|T_i, \Psi^0) = \frac{f(T_i|\theta)f(\theta)}{\displaystyle\int f(T_i|\theta')f(\theta')d\theta'} \tag{B.3}$$

Given the integral in the surrogate function $Q(\Psi|\Psi^0)$, it must be simulated by drawing R values of θ from $f(\theta|\Psi^0)$ and approximating $h(\theta|B_i, P_i, \Psi^0)$ by the weight:

$$w_{ir}^0 = \frac{L(T_i|\theta_{ir}^0)}{\displaystyle\sum_{r=1}^{R} L(T_i|\theta_{ir}^0)} \tag{B.4}$$

The lower bound function is now:

$$Q(\Psi|\Psi^0) = \sum_{i=1}^{n}\sum_{r=1}^{R} w_{ir}^0 ln(L(T_i|\theta_{ir}^0)f(\theta_{ir}^0)) \tag{B.5}$$

Maximizing this function gives a new set of parameters, Ψ^1, that guarantee $\mathcal{L}(\Psi^1) > \mathcal{L}(\Psi^0)$. Replacing Ψ^1 with Ψ^0 and iterating this process until the parameters converge. Let m denote the mth iteration of the algorithm. The parameter updates at the mth iteration are found by:

$$\hat{\tau}_{ir} = w_{ir}^m \frac{\tau_c^m normpdf(\theta_{ir}^m, \mu_c^m, \Omega_c^m)}{\sum_{c'=1}^{C} \tau_{c'}^m normpdf(\theta_{ir}^m, \mu_{c'}^m, \Omega_{c'}^m)} \tag{B.6}$$

$$\begin{aligned}
\tau_c^{m+1} &= \frac{\sum_{i=1}^{N}\sum_{i=1}^{R} \hat{\tau}_{ir}}{n} \\
\mu_c^{m+1} &= \frac{\sum_{i=1}^{N}\sum_{i=1}^{R} \hat{\tau}_{ir}\theta_{ir}^m}{\sum_{i=1}^{N}\sum_{i=1}^{R} \hat{\tau}_{ir}} \\
\Omega_c^{m+1} &= \frac{\sum_{i=1}^{n}\sum_{i=1}^{R} \hat{\tau}_{ir}(\theta_{ir}^m)(\theta_{ir}^m)'}{\sum_{i=1}^{n}\sum_{i=1}^{R} \hat{\tau}_{ir}} - (\mu_c^{m+1})(\mu_c^{m+1})'
\end{aligned} \tag{B.7}$$

Since θ are treated as observed variables, the updated parameters $\{\alpha, \beta, \sigma\}$ can be estimated by standard OLS and logit models for the continuous and binary dependent variables respectively with the weights. In particular, for simplicity, let y_i be dependent variables - the observed measures, and x_i be independent variables - unobserved factors. Equation system 3 and 5, which I want to estimate, take the form $y_i = x_i{}'\beta + u_i$.

If y_i is continuous, then

$$\begin{aligned}
&\beta^{m+1} = (XX)^{-1} * XY \\
&where\ XX = \sum_{i=1}^{N}\sum_{i=1}^{R} w_{ir}^m (x_{ir})(x_{ir})' \\
&and\ XY = \textstyle\sum_{i=1}^{N}\sum_{i=1}^{R} w_{ir}^m (x_{ir})(y_i)
\end{aligned} \tag{B.8}$$

If y_i is binary, then

$$\begin{aligned}
&\beta^{m+1} = \beta^m - B^{-1} * XY \\
&where\ B = -\tfrac{1}{4}\sum_{i=1}^{N}\sum_{i=1}^{R} w_{ir}^m (x_{ir})(x_{ir})' \\
&and\ XY = \sum_{i=1}^{N}\sum_{i=1}^{R} w_{ir}^m (x_{ir})(y_i{}' - p_{ir}^{m\prime}) \\
&with\ p_{ir}^m = \frac{(exp(x_{ir}{}'\beta))^{D_i}}{1 + exp(x_{ir}{}'\beta)}
\end{aligned} \tag{B.9}$$

I use the parameters of the measurement system, Ψ, estimated above to take individual-specific draws and use these draws as observable data to estimate investment and production functions 14 and 15

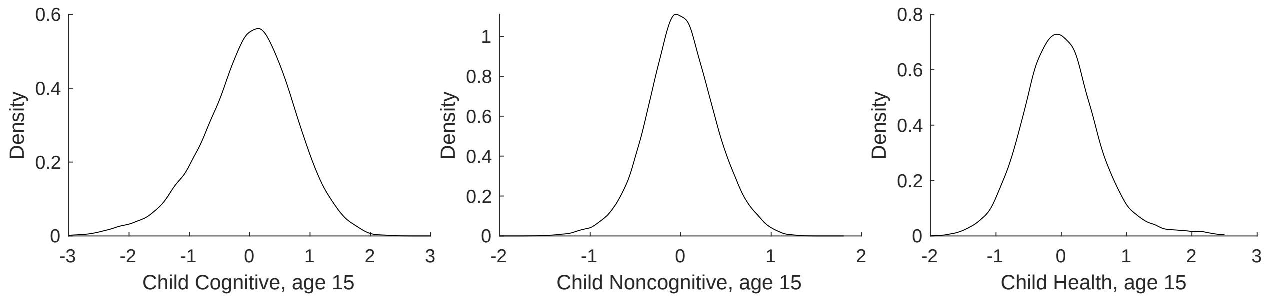
$$\begin{aligned}
lnI_{i,t} = \;& \alpha_{1,t} + \alpha_{2,t} ln\Theta^C_{i,t-1} + \alpha_{3,t} ln\Theta^{NC}_{i,t-1} + \alpha_{4,t} ln\Theta^H_{i,t-1} + \alpha_{5,t} lnP^C_i + \alpha_{6,t} lnP^{NC}_i \\
& + \alpha_{7,t} X_{i,t} + \alpha_{8,t} Z_{i,t} + v_{i,t}
\end{aligned} \tag{B.10}$$

$$ln\Theta^k_{i,t} = ln(g(\Theta^C_{i,t-1}, \Theta^{NC}_{i,t-1}, \Theta^H_{i,t-1}, I_{i,t}, P^C_i, P^{NC}_i)) + X'_{i,t}\delta^k_t + A^k_t + \mu^k v_{i,t} + \varepsilon^k_{i,t} \tag{B.11}$$

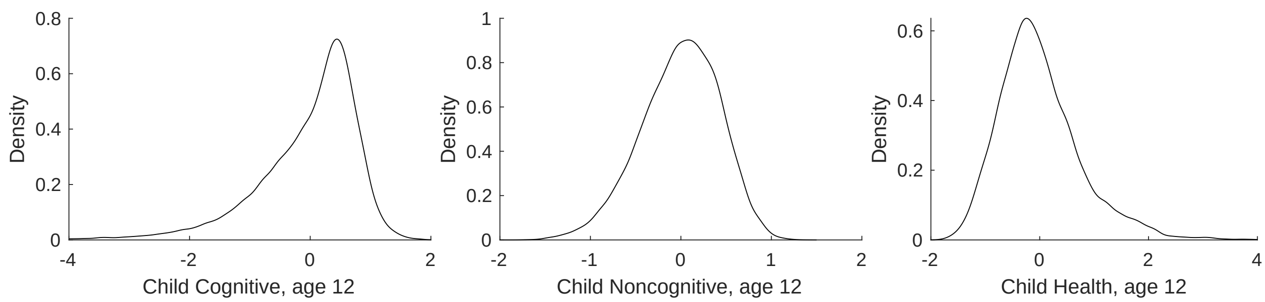
Appendix C: Factor Moments

Figure C.1: Factor Distributions

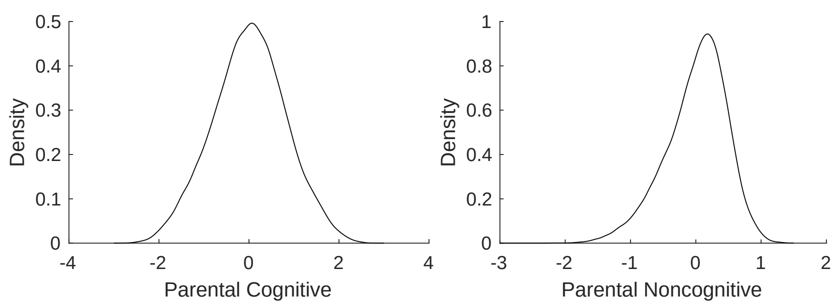
(a)



(b)



(c)



(d)

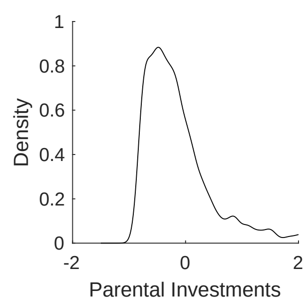


Table C.1: Factor Means, Standard Deviations and Correlations

	Child Cognitive skill, age 15	Child Noncognitive skill, age 15	Child Health, age 15	Parental Investment	Child Cognitive skill, age 12	Child Noncognitive skill, age 12	Child Health, age 12	Parental Cognitive skill	Parental Noncognitive skill
Factore means	-0.016	0.001	0.010	0.010	-0.033	-0.008	-0.001	0.007	0.001
	(0.019)	(0.010)	(0.015)	(0.017)	(0.017)	(0.009)	(0.018)	(0.018)	(0.009)
Factor standard deviations	0.765	0.381	0.623	1.001	0.834	0.435	0.758	0.818	0.471
	(0.021)	(0.015)	(0.021)	(0.050)	(0.035)	(0.013)	(0.017)	(0.015)	(0.021)
Factor correlation:									
Child's cognitive skill at age 15	1	–	–	–	–	–	–	–	–
Child's noncognitive skill at age 15	0.288	1	–	–	–	–	–	–	–
	(0.022)								
Child's health at age 15	0.275	0.115	1	–	–	–	–	–	–
	(0.024)	(0.019)							
Parental investment	0.471	0.198	0.231	1	–	–	–	–	–
	(0.018)	(0.017)	(0.020)						
Child's cognitive skill at age 12	0.830	0.293	0.265	0.383	1	–	–	–	–
	(0.014)	(0.021)	(0.022)	(0.013)					
Child's noncognitive skill at age 12	0.271	0.246	0.122	0.152	0.237	1	–	–	–
	(0.026)	(0.028)	(0.024)	(0.019)	(0.027)				
Child's health at age 12	0.367	0.164	0.835	0.381	0.295	0.146	1	–	–
	(0.021)	(0.018)	(0.016)	(0.019)	(0.019)	(0.022)			
Parent's cognitive skill	0.652	0.289	0.202	0.503	0.681	0.306	0.274	1	–
	(0.015)	(0.021)	(0.023)	(0.023)	(0.013)	(0.023)	(0.022)		
Parent's noncognitive skill	0.214	0.197	0.118	0.212	0.182	0.700	0.150	0.310	1
	(0.023)	(0.029)	(0.027)	(0.013)	(0.023)	(0.028)	(0.022)	(0.022)	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process.

Table C.2: Mixture Component Means

Factor	Type 1	Type 2
Child's cognitive skill at age 15	-0.209	0.608
	(0.028)	(0.032)
Child's noncognitive skill at age 15	-0.037	0.123
	(0.012)	(0.018)
Child's health at age 15	-0.107	0.386
	(0.018)	(0.033)
Parental investments	-0.324	1.085
	(0.014)	(0.097)
Child's cognitive skill at age 12	-0.201	0.511
	(0.029)	(0.016)
Child's noncognitive skill at age 12	-0.030	0.062
	(0.011)	(0.029)
Child's health at age 12	-0.223	0.714
	(0.020)	(0.055)
Parental cognitive skill	-0.181	0.611
	(0.025)	(0.048)
Parental noncognitive skill	-0.062	0.202
	(0.012)	(0.026)
Type share	0.763	0.237
	(0.019)	(0.019)

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process.

Appendix D: Marginal Products of the CES Functions

This appendix derives the marginal products of inputs for the CES production function indicated in Equation 2, which is:

$$\begin{aligned}\Theta_{i,t}^{k} = \ & \left[\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}} + \gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}} + \gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}} + +\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}\right. \\ & \left.\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}} + \gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}\right]^{1/\rho^{tk}} e^{X_{it}'\delta_{t}^{k}+A_{t}^{k}+\varepsilon_{it}^{k}}\end{aligned} \tag{D.1}$$

Then the marginal product of cognitive skills at time t (age 12) with the outputs being cognitive skills, noncognitive skills and health at time t (age 15) - $ln(\Theta_{i,t}^{k})/ln(\Theta_{i,t-1}^{C})$ for k $\in$ (C, N, H) are derived as follows:

$$\begin{aligned}ln(\Theta_{i,t}^{k}) = \ & \tfrac{1}{\rho^{tk}} ln\left[\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}} + \gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}} + \gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}} + \gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}} +\right. \\ & \left.\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}} + \gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}\right] + X_{it}'\delta_{t}^{k} + A_{t}^{k} + \varepsilon_{it}^{k}\end{aligned} \tag{D.2}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial \Theta_{i,t-1}^{C}} = \frac{1}{\rho^{tk}}\left(\frac{1}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right)\rho^{tk}\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}-1}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial \Theta_{i,t-1}^{C}/\Theta_{i,t-1}^{C}} = \frac{1}{\rho^{tk}}\left(\frac{1}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right)\rho^{tk}\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{C}} = \left(\frac{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.3}$$

Similarly, the marginal products of child noncognitive skills, child health, parental investment and parental cognitive and noncognitive skills are:

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{NC}} = \left(\frac{\gamma_{2,t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.4}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{H}} = \left(\frac{\gamma_{3,t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.5}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{I}} = \left(\frac{\gamma_{4,t}^{k}(\Theta_{i,t-1}^{I})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.6}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{PC}} = \left(\frac{\gamma_{5,t}^{k}(\Theta_{i,t-1}^{PC})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.7}$$

$$\frac{\partial ln(\Theta_{i,t}^{k})}{\partial ln\Theta_{i,t-1}^{PN}}=\left(\frac{\gamma_{6,t}^{k}(\Theta_{i,t-1}^{PN})^{\rho^{tk}}}{\gamma_{1,t}^{k}(\Theta_{i,t-1}^{C})^{\rho^{tk}}+\gamma_{2t}^{k}(\Theta_{i,t-1}^{NC})^{\rho^{tk}}+\gamma_{3t}^{k}(\Theta_{i,t-1}^{H})^{\rho^{tk}}+\gamma_{4t}^{k}(I_{i,t})^{\rho^{tk}}+\gamma_{5t}^{k}(P_{i}^{C})^{\rho^{tk}}+\gamma_{6t}^{k}(P_{i}^{NC})^{\rho^{tk}}}\right) \tag{D.8}$$

Appendix E: Estimates of Production Functions without Endogenous Investments

Table E.1: Estimates of Production Functions without Endogenous Investments

	Cognitive skills at age 15 (1)	Noncognitive skills at age 15 (2)	Health at age 15 (3)
Child's cognitive skills at age 12	0.619***	0.000	0.035
	(0.044)	(0.000)	(0.028)
Child's noncognitive skills at age 12	0.152***	0.425***	0.123***
	(0.055)	(0.080)	(0.043)
Child's health at age 12	0.079***	0.071***	0.801***
	(0.020)	(0.021)	(0.025)
Parental investments	0.120***	0.396***	-0.051**
	(0.031)	(0.026)	(0.021)
Parental cognitive skills	0.071**	0.007	0.012
	(0.033)	(0.006)	(0.030)
Parental noncognitive skills	-0.042	0.101	0.080*
	(0.047)	(0.074)	(0.044)
A_t	-0.033	0.223***	0.164***
	(0.021)	(0.020)	(0.021)
Female	0.071***	-0.027*	-0.212***
	(0.015)	(0.014)	(0.023)
Urban	0.071***	-0.259***	-0.129***
	(0.027)	(0.025)	(0.022)
Number of siblings aged 0-18	0.008	0.035***	0.002
	(0.008)	(0.007)	(0.006)
Complementarity(ρ)	-0.150	-2.423***	-0.355
	(0.137)	(0.259)	(0.224)
Elasticity of substitution	0.870***	0.292***	0.738***
	(0.122)	(0.023)	(0.111)
Observations		*961*	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.

Table E.2: Marginal Effects - Production Functions without Endogenous Investments

	Cognitive skills at age 15 (1)	Noncognitive skills at age 15 (2)	Health at age 15 (3)
Child's cognitive skills at age 12	0.619***	0.001	0.036
	(0.043)	(0.001)	(0.028)
Child's noncognitive skills at age 12	0.152***	0.356***	0.125***
	(0.055)	(0.061)	(0.043)
Child's health at age 12	0.079***	0.107***	0.796***
	(0.019)	(0.019)	(0.022)
Parental investments	0.120***	0.427***	-0.052**
	(0.032)	(0.024)	(0.022)
Parental cognitive skills	0.071**	0.015*	0.012
	(0.033)	(0.009)	(0.032)
Parental noncognitive skills	-0.042	0.094	0.082*
	(0.048)	(0.063)	(0.046)
Observations		*961*	

Note: Standard errors in parentheses based on 100 bootstrap replications of the entire estimation process; *** p<0.01, ** p<0.05, * p<0.1.