UBS International Center of Economics in Society

at the University of Zurich

University of Zurich[UZH]

UBS Center Working Paper Series
Working Paper No. 17, November 2016

ISSN 2296-2778 (Online)

Health and Skill Formation in Early Childhood

Pietro Biroli

About the UBS International Center of Economics in Society

The UBS International Center of Economics in Society, UBS Center in short, is an Associated Institute at the Department of Economics of the University of Zurich. It was established in 2012, enabled by a founding donation by UBS, which the bank made on the occasion of its 150th anniversary. In view of the very generous donation, the university named the UBS Center after its benefactor. The new center of excellence serves two main aims, the first of which is to finance world-class research in economics on all levels, to be conducted at the University's Department of Economics. It thereby supports the department's ambition to further strengthen its position as one of the top economics departments in Europe and to make Zurich one of the best places for research in economics in the world. The Center's other aim is to serve as a platform for dialogue between academia, business, politics, and the broader public, fostering continuous knowledge transfer in the process. Delivering on these aims will in turn also contribute to the quality of Zurich, and Switzerland more generally, as leading locations for education and business.

Health and Skill Formation in Early Childhood

Pietro Biroli
University of Zurich,
Department of Economics
and UBS Center

Abstract

This paper analyzes the developmental origins and the evolution of health, cognitive, and socio-emotional skills during early childhood, from age 0 to 5. We explicitly model the dynamic interactions of health with the child's behavior and cognitive skills, as well as the role of parental investment. A dynamic factor model corrects for the presence of measurement error in the proxy for the latent traits. Using data from the Avon Longitudinal Study of Parents and Children (ALSPAC), we find that children's capabilities strongly interact and build on each other: health is an important determinant of early socio-emotional development; in turn socio-emotional skills have a positive impact on the evolution of both health and cognitive functions; on the other side, the effect of cognitive abilities on health is negligible. Furthermore, all facets of human capital display a high degree of persistence. Finally, mother's investments are an important determinant of the child's health, cognitive, and socio-emotional development early in life.

Keywords: Human capital, health, early childhood, family investment, intergenerational transmission, ALSPAC.
JEL classification: J24, J13, I10, I12, I14.

Acknowledgements
I benefited from helpful comments from Dan Black, Hoyt Bleakley, Flavio Cunha, George Davey-Smith, Miriam Gensowski, Tim Kautz, Robert Lalonde, Maria Rosales-Rueda, Daniel Tannenbaum, Frank Windmeijer, and the participants at various workshops. I am especially grateful to Gabriella Conti, Steven N. Durlauf, and James J. Heckman for their continued support of this research.

Contact
Pietro Biroli, Department of Economics and UBS International Center of Economics in Society, University of Zurich, Switzerland. E-mail: pietro.biroli@econ.uzh.ch

I. Introduction

Harry Truman once said "A nation is only as healthy as its children." Indeed, various disciplines have accumulated evidence on the fundamental role played by early childhood health in shaping wellbeing later in life.[1] At the same time, an emerging developmental literature has demonstrated the importance of early cognitive and socio-emotional skills[2]: intelligence and cognition are a main ingredient in economic success and personal wellbeing;[3] the parallel importance of socio-emotional abilities and personality in influencing later life outcomes has been extensively studied by psychologist and, more recently, economists.[4] Overall, the fundamental role that skills and capabilities[5] play in achieving a long and successful life has long been recognized, and the subsequent importance of investing and developing such abilities is widely acknowledged.[6]

However most studies focus on a narrowly defined set of capabilities and usually fail to recognize or to properly estimate the rich set of complementarities and interconnections among different skills. The main contribution of this paper is to undertake a comprehensive approach that integrates health in a unifying framework of human capital formation, considered as a multidimensional asset that dynamically evolves in the family environment. We achieve this goal by considering a simple economic model where future skills are generated by combining the past stock of the child's human capital, various parental abilities, and different types of investment. This structure enables us to estimate the degree of self- and cross-productivity of skills, and we are able to evaluate the degree of intergenerational transmission of human capital and compare it to the effectiveness of investments in the realm of parenting, curative health care, and preventive health care.

We formulate and estimate the model by building on various strands of literature. We use extensively the structure developed by Cunha and Heckman (2007, 2008) and generalized in Cunha et al. (2010), and integrate it with the seminal model of health formation by Grossman (1972), extending it to the early childhood period. Like Grossman, we consider the relation between health and investment choices; however, he considers the health endowment and the preferences of adults as exogenous to his model, while our analysis can shed light on how these important initial inputs are formed.

Previous work by Palloni et al. (2009), Conti et al. (2010), and Conti and Heckman (2010) evaluates the joint effects of health, cognitive skills, and socio-emotional ability on adult outcomes. However, they do not have detailed data on the *evolution* of health over childhood, and therefore they focus on the long-term relation between

[1] See the seminal epidemiological work of Barker et al. (1989b); Barker et al. (1989a); more recent work by Gluckman and Hanson (2006); Case et al. (2005); Smith (2009); Goodman et al. (2011); as well as literature reviews of Currie (2009); Currie et al. (2010); Bleakley (2010); Currie and Almond (2011).

[2] See Cunha et al. (2006); Heckman (2007); Cunha and Heckman (2007); Almlund et al. (2011)

[3] See, among many others, Cawley et al. (2001) Jokela et al. (2009)

[4] See Heckman et al. (2006); Borghans et al. (2008); Almlund et al. (2011); Hampson (2012); Heckman (2012); Kautz et al. (2014); Heckman and Kautz (2012). Several terms have been used to refer to these skills: noncognitive, socio-emotional, 'soft' skills, etc. This paper uses the term 'socio-emotional' skills throughtout.

[5] Interestingly, each field of study uses different terms to refer to similar underlying concepts: skills, abilities, character, personality, aptitudes, traits, human capital, capabilities, and so on. The purpose of this paper is not to spur a philosophical debate to highlight the distinctions between these terms, but rather to take an empirical approach that considers only a core set of broadly defined latent constructs. Therefore we will use these terms interchangeably throughout the paper. See Sen (1990, 1985) and Nussbaum (2011) for a general discussion of 'functionings' and 'human capabilities'.

[6] "The most valuable of all capital is that invested in human beings; and of that capital the most precious part is the result of the care and influence of the mother" *Marshall (1890), paragraph VI.IV.11*

later life outcomes and health endowments measured in one period early in life. Following the early contribution of Shakotko et al. (1980b) and Shakotko et al. (1980a), we focus on the evolution of health during the early childhood period, and its impact on human capital formation.[7]

Furthermore, we contribute to the literature on environmental determinants of health and human capital of the child. A lot of emphasis has been given to socioeconomic determinants of health, especially by Marmot (2010); Duncan and Magnuson (2005); Hertzman and Boyce (2010); however there is strong evidence that also parenting factors,[8] stress and maltreatment,[9] and Early Childhood Interventions[10] play a very important role in shaping adult health and wellbeing. We contribute to this debate by evaluating the relative importance of different dimensions of parental investment and compare them to the effect of maternal characteristics on the development of the child.

In our empirical analysis we use data from a prospective cohort of British children, followed since birth. We find substantial evidence of self-productivity and a strong persistence of human capital: early insults to health, cognitive abilities and socio-emotional skills have a long lasting impact over childhood. Furthermore, the characteristics of the mother play an important role, especially early in life. We find that family environment and parental investment have a significant effect on childhood development, especially in the areas of cognitive and socio-emotional abilities, while curative care and preventive investment in health have an important influence on the child's physical health and socio-emotional wellbeing. Furthermore, the skills of the mother strongly relate to the ones of the child: her cognitive ability have a prominent effect in the cognitive development of the child; her health has a direct impact on the health of her offspring; her socio-emotional skills determine the socio-emotional skills of the kid. In other words, we show evidence supporting the intergenerational transmission of capabilities. Finally, we find that health is a fundamental dimension of human capital formation. While it does not seem to be strongly related to cognitive development, it displays an important degree of cross-linkage with socio-emotional development.

The remainder of the paper is structured as follows. Section 2 briefly introduces the model, outlines the estimation strategy and then discusses the data and the measurement used throughout the paper. Section 3 presents the estimation results and comments on the main findings. Section 4 concludes.

II. Health Formation and Development of Skills

A. The Model

Our main interest lies in investigating the dynamic evolution of human capital during childhood. A simple conceptual framework that suits the purpose was proposed by Cunha and Heckman (2007) and we will follow it throughout the paper.

[7]There is a growing literature that analyzes the evolution of the health-income gradient as the child ages, see Case et al. (2002); Currie and Stabile (2003); Currie et al. (2007); Chatterji et al. (2012); however they focus on the relation between income and health at different ages, rather than on the evolution of child health over time. Condliffe and Link (2008) lay down a dynamic model of health development over childhood, but they do not estimate it because of lack of data.

[8]See Case and Paxson (2002); Stewart-Brown et al. (2005); Belsky et al. (2007); Waylen et al. (2008).

[9]See Danese et al. (2011, 2009).

[10]See Duggan et al. (1999); Olds (2002); Muennig et al. (2009); Heckman et al. (2010).

Consider multiple periods of childhood $t \in 1, 2, ..., T$, $T \geq 2$, followed by τ periods of adult working life; the stock of human capital of a child in period t is represented by a multivalued vector $\boldsymbol{\theta}_t$. We assume that at each period the human capital of the child can be decomposed into three broad categories: cognitive abilities, θ_t^C, socio-emotional skills, θ_t^{SE} and health, θ_t^H.[11]

The evolution of human capital over childhood has many determinants: the past stock of the child's skill, $\boldsymbol{\theta}_{t-1}$, the stock of parental ability, $\boldsymbol{\theta}_P$, as well as the quality and amount of care, time and goods that the family invests into the development of the child, $\boldsymbol{\theta}_t^I$. Each of these inputs is potentially a multivalued vector with different components. Regarding parental abilities, we follow a similar approach as for the human capital of the child and consider three types of parental skills[12]: cognitive ability, θ_P^C, socio-emotional skills, θ_P^{SE} as well as physical health, θ_P^H. As far as the investment dimension is concerned, in this paper we consider a factor that captures parenting decisions and practices, θ_t^{PI}, a factor of preventive investment in the health of the child, θ_t^{PH}, as well as the dimension of curative health care, θ_t^{CH}.

Finally, the specification of an initial condition closes the model: assume that each agent is endowed at birth with θ_0, which captures both the family environment and other factors that have a direct influence on birth conditions.[13] Similarly to parental characteristics, we consider the effect that this initial endowment has on the evolution of human capital.

Combining everything together, we have the following general form for the production function of next period human capital:

$$\boldsymbol{\theta}_{t+1} = \mathbf{f}\left(\boldsymbol{\theta}_t, \boldsymbol{\theta}_t^I, \boldsymbol{\theta}_P, \theta_0, \eta_t\right) \tag{1}$$

where η_t captures unobserved inputs and shocks that affect the accumulation of skills, and we assume that $f^k(\cdot)$ is monotone increasing in its arguments and twice continuously differentiable for $k \in \{C, SE, H\}$.

One of the problems that we have to face is that there is no natural scale to pin down the distribution of $\boldsymbol{\theta}$. In order to overcome this drawback, we rely on the idea that the ultimate goal of investing in human capital development is to enable the child to live a long and successful life. What we care about are not mother's reports on temperament or wellbeing, but rather important lifetime outcomes such as the ability to avoid sickness, criminal activity, teen pregnancy or drug abuse. In order to capture this important aspect of skill formation, we analyze how adult outcomes Q_j, $j \in \{1, ..., J\}$ originate from the various elements of the final stock of child's human capital $\boldsymbol{\theta}_{T+1}$:

$$Q_j = g_j\left(\theta_{T+1}^C, \theta_{T+1}^{SE}, \theta_{T+1}^H\right) \tag{2}$$

By doing so, the evolution of each facet of human capital can be related to its productivity in the achievement of Q_j. Furthermore, considering multiple outcomes

[11] See Appendix B.B2 for a discussion of the assumption to use three categories.

[12] Theoretically also parental characteristics could change and evolve over time; since the evolution of skills in adulthood is not the main focus of this paper, for simplicity we will consider parental characteristics to be time constant so that $\theta_{P,t}^k = \theta_P^k$ for all t and k

[13] See for example Olds (2002) and Levitt (2003), as well as the extensive research of the effect of mother's choices during pregnancy on birth weight, Currie and Moretti (2007)

allows for a richer characterization of the impact of each human skill; for instance, child health could be more relevant in determining sickness later in life, while child socio-emotional skills might be more relevant for adult depression.[14]

It is worth noticing that the formulation used in (2) assumes that only the final stock of human capital $\boldsymbol{\theta}_{T+1}$ is relevant in explaining the outcomes Q_j, which are not allowed to depend on previous skills $\boldsymbol{\theta}_{T-t}$ or the evolution of such skills. In the empirical part we tested this assumption by including additional lags to equation (2) and we did not find a substantial difference in the main results.

B. The Estimation strategy

The main issue with estimating technology (1) is that both inputs and outputs can only be proxied: we do not observe the cognitive ability of a child, or the status of his health. However there are many observable measures that are a manifest expression of these latent traits, such as the result of a test or the number of times that a child was sick; each of them can be measured more easily, albeit with a degree of error, and used as an instrument to trace the distribution of the unobserved variables of interest.

This problem has been extensively addressed in the field of psychometrics. Following Carneiro et al. (2003) and Hansen et al. (2004), we use a linear measurement system to identify the joint distribution of the latent factors θ.[15] Specifically, we have access to multiple measurements $\left\{M_t^{k,l}\right\}_{l=1}^{L_t^k}$ that we assume to be dedicated to a particular latent factor $k \in \{C, SE, H, PI, PH, CH, P\}$.[16] In this notation, L_t^k denotes the number of measurements available at time t for each factor k. Assuming a linear dependence between the measurement and the factors, we have the following system of equations:

$$
\begin{aligned}
M_t^{C,1} &= \phi_t^{C,1} X_t + \alpha_t^{C,1} \theta_t^C + \varepsilon_t^{C,1} \\
&\vdots \\
M_t^{C,L_t^c} &= \phi_t^{C,L_t^c} X_t + \alpha_t^{C,L_t^c} \theta_t^C + \varepsilon_t^{C,L_t^c} \\
M_t^{SE,1} &= \phi_t^{SE,1} X_t + \alpha_t^{SE,1} \theta_t^{SE} + \varepsilon_t^{SE,1} \\
&\vdots \\
M_t^{SE,L_t^{SE}} &= \phi_t^{SE,L_t^{SE}} X_t + \alpha_t^{SE,L_t^{SE}} \theta_t^{SE} + \varepsilon_t^{SE,L_t^{SE}} \\
M_t^{H,1} &= \phi_t^{H,1} X_t + \alpha_t^{H,1} \theta_t^H + \varepsilon_t^{H,1} \\
&\vdots \\
M_t^{H,L_t^H} &= \phi_t^{H,L_t^H} X_t + \alpha_t^{H,L_t^H} \theta_t^H + \varepsilon_t^{H,L_t^H}
\end{aligned}
\tag{3}
$$

or in compact form $\mathbf{M}_t = \boldsymbol{\Phi}_t \mathbf{X}_t + \mathbf{A}_t \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t$.

[14] See Cunha et al. (2006) for a review of such evidence

[15] See also Williams (2013) for a discussion of non-parametric estimation with discrete measures and continuous latent variables.

[16] Dedicated refers to the assumption that each measure is only related to one latent variable, so that for example 'having a stomach ache' is only related to health θ^H and not to cognition θ^C. Relaxing this assumption awfully increases the number of parameters to be estimated $\left[\text{by } (k-1) \times \sum_t \sum_k L_t^k\right]$ without yielding significant improvements to the estimation. Exploratory Factor Analysis (EFA) on the measurements used in this paper provides support for this assumption, since the factor loadings $\alpha_t^{k,L}$ for measurements other than the dedicated ones are usually very close to zero.

$\phi_t^{kl} X_t$ is the mean of proxy l at time t for factor k, which can depend from some observable characteristics X; $\varepsilon_t^{k,l}$ captures the noise inherent in that measurement. Both the mean and the error term are assumed to be independent of the latent trait θ_t^k.

Assuming that the error terms $\varepsilon_t^{k,l}$ are orthogonal to each other and over time, and using the related restrictions on the variance-covariance matrix of the measurements, we can identify the factor loadings $\alpha_t^{k,l}$ and the variance of the latent factors up to a scale.[17] In order to achieve point identification of all the parameters, we need a normalization. As for common practice, we normalize the first loading of each factor to be one: $\alpha_t^{k,1} = 1$ for all t and k.[18] Finally, following Schennach (2004), we can identify the distribution of θ_t^k and the error ε_t^k. In other words, the identification comes from the idea that the latent trait θ_t^k is the only component that drives the common covariance of the measurements, once the effect of the observable X has been netted out, while the rest of the variation is due to the noise.[19]

Assuming dedicated measurements or independent errors is not necessary for identification and relaxing these assumptions does not yield substantial difference in the results.[20] It is worth stressing that these assumptions do not require measurements of different facets of human capital to be uncorrelated (this is not true in the data); rather, the underlying assumption is that the correlations that we observe in the data are only driven by the correlation of the latent factors.[21] The same approach leads to the identification of investment $\boldsymbol{\theta}_t^I$ and parental abilities $\boldsymbol{\theta}_P$.

Once the joint distribution of the various inputs and outputs $\left\{\boldsymbol{\theta}_t, \boldsymbol{\theta}_t^I, \boldsymbol{\theta}_P\right\}$ is identified, it is possible to identify the relevant parameters of the technology (1) using the conditional correlation between the estimated factors. For simplicity, we follow Cunha and Heckman (2008) and estimate a linear specification of the technology function $\mathbf{f_t}(\cdot)$.Therefore we estimate the following system of linear equations:

$$\begin{aligned}\begin{pmatrix} \theta_{t+1}^{SE} \\ \theta_{t+1}^{C} \\ \theta_{t+1}^{H} \end{pmatrix} &= \begin{pmatrix} \gamma_1^{SE} & \gamma_2^{SE} & \gamma_3^{SE} \\ \gamma_1^{C} & \gamma_2^{C} & \gamma_3^{C} \\ \gamma_1^{H} & \gamma_2^{H} & \gamma_3^{H} \end{pmatrix} \begin{pmatrix} \theta_t^{SE} \\ \theta_t^{C} \\ \theta_t^{H} \end{pmatrix} + \begin{pmatrix} \delta_1^{SE} & \delta_2^{SE} & \delta_3^{SE} \\ \delta_1^{C} & \delta_2^{C} & \delta_3^{C} \\ \delta_1^{H} & \delta_2^{H} & \delta_3^{H} \end{pmatrix} \begin{pmatrix} \theta_t^{PI} \\ \theta_t^{PH} \\ \theta_t^{CH} \end{pmatrix} \\ &+ \begin{pmatrix} \beta_1^{SE} & \beta_2^{SE} & \beta_3^{SE} \\ \beta_1^{H} & \beta_2^{H} & \beta_3^{H} \\ \beta_1^{0} & \beta_2^{0} & \beta_3^{0} \end{pmatrix} \begin{pmatrix} \theta_P^{SE} \\ \theta_P^{H} \\ \theta_0 \end{pmatrix} + \begin{pmatrix} \eta_t^{SE} \\ \eta_t^{C} \\ \eta_t^{H} \end{pmatrix}\end{aligned}$$

Assuming joint normality of the error terms $\boldsymbol{\varepsilon}_t, \eta_t$ we can jointly estimate the measurement equations (3) and the technology equation (1) using a Maximum Likelihood Estimator.[22]

Another possible way of looking at the data is to solve the above system back-

[17]Identification of the model can be achieved using less stringent restrictions on the variance-covariance matrix of the errors, see Cunha et al. (2010) for a more thorough discussion on non-classical measurement error.

[18]Of course, the choice of which measurement is considered to be "the first" is discretionary. For each outcome we purposefully choose a normalization method so that a higher value of the latent factor corresponds to a more favorable outcome. We choose the first measure to be the one that (usually) has the highest factor loading, but we also try to be consistent over time periods so that the same measurement is used for normalization. See appendix B.B1 for more details

[19]Notice that this identification technique is very similar to an instrumental variable approach, where each measure is used as a proxy for the latent variable

[20]See appendix B for more details on the robustness checks; for a discussion of general assumptions on the matrix of factor loadings $\mathbf{A}$ needed for identification of factor models see Anderson and Rubin (1956) and Lopes and Fruhwirth-Schnatter (2010)

[21]See Tables A1 to A5 for the estimated correlation among the factor scores.

[22]For alternative methods of estimation, see Appendix C

wards, expressing the final stock of skills at age $T+1$ as a function of all the series of previous investments as well as the initial condition and parental abilities. This technique allows for a more straightforward characterization of the pattern of parental investment and its relative importance, but it is not able to capture the potential for complementarities across different facets of human capital.

C. The Data

One of the strengths of this analysis is the use of very precise measurements from the Avon Longitudinal Survey of Parents and Children (ALSPAC), an extremely rich dataset collected by epidemiologists at the University of Bristol. The ALSPAC follows prospectively a cohort of children born from mothers living in a health district in the former County of Avon, in the South West of England, with an expected delivery date between April 1991 and December 1992.[23] The children from 14,541 pregnancies were initially recruited. For this analysis we excluded all multiple births, children with major congenital malformations or illnesses, those who did not survive or dropped out of the study before 12 months, as well as children of minorities and armed forces, leaving a cohort of 11,948 infants.[24]

Detailed information has been collected since pregnancy using self-administered questionnaires, data extraction from research clinics and medical notes, linkage to routine information systems. This dataset contains state-of-the-art measures of cognition, behavior, and health of the child as well as indicators of personality, physical and mental wellbeing of both the mother and the father (when present). An extensive report of the relationship between the caregiver and the child as well as the choice of parenting practices allows for a precise characterization of the decision to invest in the capabilities of the child. Furthermore, detailed reports of the child's food intake, activity level and utilization of medicine and health care facilities has been collected. Coupled with the fact that the behavior of the mother has been followed as early as 8 weeks into pregnancy, this enables a precise analysis of the evolution of investment over the first part of the life-cycle, capturing the dynamics that shape the evolution of health from early in the womb all the way to late childhood.[25]

This great wealth of data enabled us to obtain all the measurements $M_t^{k,l}$ necessary for the estimation of the model. One of the contributions of the paper is to focus on common variations in child health, measured using mothers assessment of the child's health and wellbeing,[26] coupled with reports of specific illnesses that are very common in infant and toddlers (from cough and temperature to stomach ache). In other words, health is measured using common symptoms and acute illnesses, rather than complex and chronic health conditions. Therefore, instead of focusing on a minority of very sick children, we consider the normal variation in health that virtually every child in our dataset experiences.[27] Regarding the other facets of human capital,

[23] See Fraser et al. (2013); Boyd et al. (2013)

[24] 14,541 is the initial number of pregnancies for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a "Children in Focus" clinical visit by 19/07/99. Of these initial pregnancies, there was a total of 14,676 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age.

[25] Please note that the study website contains details of all the data that is available through a fully searchable data dictionary, `http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/`. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

[26] Reported using a Likert scale from 1-"Very healthy, no problems" to 4-"Almost always unwell"

[27] In our sample, 99.7% of the children experience at least one of the symptoms that we use as measures

cognitive ability of the child is evaluated with the MacArthur Infant Communication questionnaire (Fenson et al. (1991)), as well as a revised version of the Denver Developmental Screening Test (Frankenburg and Dodds (1967)), which relate to 4 different categories: social and communication skills, fine motor skills, hearing and speech, gross motor skills. Socio-emotional skills were elicited using the Carey Temperament Scale for the first three years of life (Carey and McDevitt (1977), Fullard et al. (1978)), and the Revised Rutter Parent Scale for Preschool Children until age 5 (Elander and Rutter (1996)). The extent of parental involvement in the development of the child was assessed using questions adapted from the HOME Inventory (Caldwell and Bradley (1984)), while the parental intervention on the realm of child health was measured using questions regarding feeding behavior, frequency of visits or calls to the doctor and usage of medicines.

Notice that measurements need not be the same across different ages: as the child evolves, so do her skills and capabilities; therefore also the scales and the battery of tests used should be age-appropriate and change in order to better reflect these developments. For example, it would be preposterous to ask a one-year-old whether she can read full sentences or count up to 100, which instead is asked in the Denver Test for 5-year-olds. Similarly, mothers were asked whether they took their child to the library only after age 2. On the other side, some questions are relevant for all ages and should be asked consistently throughout the time periods, as is the case for the questions about feeding or visits to the doctor. This aspect of flourishing and maturation is reflected very well in the dataset that we use: as the child ages, some questions are always asked to the respondent while others change in order to capture the relevant aspects of child development and the richness of family environment.

Finally, mother abilities were evaluated at baseline before the birth of the child, in order to obtain a stable measure that was not influenced by the current relation with the child but would rather reflect the long-run characteristics of the care giver.[28] Socio-emotional skills were accounted for by the Inter Personal Sensitivity Measure (IPSM, Boyce and Parker (1989)), while measures of overall wellbeing and chronic health conditions were used to evaluate her physical health. Since no intelligence test or similar questions were asked to the mother, we use the highest grade achieved as a proxy for her intellectual ability. Although not a perfect measure of θ_P^C, education is strongly related to cognitive ability and it is a variable of interest on its own, since very easily available in most datasets and widely studied in the literature of intergenerational transmission of wealth and capabilities.[29]

Tables $(1-4)$ provide more details on the measurements we used for the different time periods.

to construct the health factor. On average, every time period children display 2 to 3 of these common illnesses. On the other hand, in a cross-country comparison Merrick and Carmeli (2003) estimate childhood disabilities raging from 5.8% in the US to 9.8% in Finland.

[28]This is important especially for socio-emotional skills and health: we do not want to mistakenly take into account a transitory change in mood or health due to pregnancy or tense relationship with the child

[29]See for example the seminal paper Heath et al. (1985), as well as the discussion in Bowles and Gintis (2002) and the analysis in Carneiro et al. (2013)

Table 1: Measurements used for Cognitive skills, Socio-emotional skills and Health

t = 1 **0 to 4 months**	**t = 2** **6m to 1y3m**	**t = 3** **1y6m to 2y6m**	**t = 4** **2y7m to 3y7m**	**t = 5** **3y11m to 4y9m**
		Cognitive Skills		
Baby laughs	Vocabulary knowledge	Vocabulary knowledge	Vocabulary knowledge	Reading and counting
Baby looks at mum's face	Understanding	Language knowledge	Past tense knowledge	Playing and sharing
Baby follows mum with eyes	Non-verbal communication	Plurals knowledge	Plurals knowledge	Listening and singing
Baby smiles	Communication	Grammar	Word combination	Social skills
Baby squeals	Social skills	Communication	Social development	Drawing skills
Baby lifts head when on tummy	Social development	Social achievement	Fine motor	Building skills
Baby touches hands together	Fine motor	Social achievement		
	Gross motor	Fine motor		
		Socio-emotional skills		
Baby is grizzly	Adaptability	Adaptability	Conduct difficulties	Cries and is fussy
Baby is placid	Approach	Approach	Emotional difficulties	Activity
Baby is fretful	Mood	Mood	Hyperactivity	Aggressivity
Baby is demanding	Rhythmicity	Rhythmicity	Prosocial	Relation with other children
Baby is angry	Persistence	Persistence		Sociability
Baby is stubborn	Distractibility	Activity		Concentration
Baby is happy		Intensity		
		Health		
Health of child in past month	Health of child in past month	Health of child in past month	Health of child in past month	Health of child in past month
Baby vomits	Health of child in 1st months		Health of child in past year	Health of child in past year
Baby possets	Child had cough	Child had cough	Child had cough	Child had cough
Baby ever had .. cough	... diarrhea	... diarrhea	... vomiting	... vomiting
... high temperature	... vomiting	... vomiting	... stomach ache	... stomach ache
... sniffles	... high temperature	... stomach ache	... high temperature	... high temperature
	... cold	... high temperature	... earache	... earache
		... earache	... rash	... rash
			... headache	... headache
				Number of infections

Table 2: Measurements used for Parental Investment and Health Investment

t = 1 **0 to 4 months**	**t = 2** **6m to 1y3m**	**t = 3** **1y6m to 2y6m**
	Parental investment	
Feelings about pregnancy	Mum plays games with child	Mum plays games with child
Reaction to pregnancy	Mum and child play with toys	Mum plays with toys with child
Motherhood means personal sacrifice	Mum shows child picture books	Taken to interesting places
Motherhood gives new opportunities		Taken to library
	Mum talks to child while working	Mum talks to child while working
		Mum reads to the child
	Mum tries to teach child	Mom teaches child
	No of books child owns	No of stimulating toys
	Maternal bonding	Maternal experience
	Preventive Health Investment	
Breastfeeding duration	Child being choosy with food	Child been choosy with food
Child refuses milk	Child refuses food	Child refuses the right food
Baby fed on regular schedule	Child has eating routine	Child has eating routine
Difficulties feeding	Difficulties feeding	
	Curative Health Investment	
Ever called out doctor for baby	Visits to doctor	Doctor has seen child at surgery
No. Of medications since at home	Doctor visits to home	Doctor called to home
Intention to immunize baby		No. Of doctor visits to home
	Child had cough medicine	Child had cough medicine
	Child had antibiotics	Child had antibiotics
	Child had diarrhea medicine	Child had diarrhea medicine

III. Estimates of the Technology

We now turn to the empirical estimates of our model.

A. Omitting the health factor

Let us begin with the estimation of the development of children's cognitive and socio-emotional skills, temporarily omitting the health factor and the health investments:

$$\begin{aligned}\theta_{t+1}^{k} &= \gamma_1^k \theta_t^C + \gamma_2^k \theta_t^{SE} + \delta_1^k \theta_t^{PI} \\ &\quad + \beta_1^k \theta_P^C + \beta_2^k \theta_P^{SE} + \beta_3^k \theta_0 + \eta_t^k\end{aligned}$$

for $t \in \{1, ..., 5\}$ and $k \in \{C, SE\}$. Columns (1)-(4) of table (5) show the results for the whole sample of 11,948 children, first omitting and then including the effect of the birth conditions.

There is evidence of substantial self-productivity of skills: current level of a skill have a strong relation to its future levels. Furthermore, a certain degree of cross-productivity can be found for socio-emotional skills, whose current level is related to the development of cognitive abilities. Besides the past stock of a particular facet of human capital, parental investment is the input that most strongly benefits the development of both cognitive and socio-emotional capabilities. Finally, it seem that mother's education fosters the formation of cognitive skills but has no effect on the

Table 3: Measurements used for Parental Investment and Health Investment

t = 4 **2y7m to 3y7m**	**t = 5** **3y11m to 4y9m**
Parental investment	
Mum plays games with child	Mum makes things with child
Mum plays with toys with child	Mum plays with toys with child
Taken to interesting places	Taken to interesting places
Taken to library	Taken to library
Mum talks to child while working	Mum talks to child while working
Mum reads to the child	Mum reads to the child
Mom teaches child	Mum draws or paints with child
Maternal Enjoyment	
Maternal Confidence	
Preventive Health Investment	
Child been choosy with food	Child been choosy with food
Child refuses the right food	Child refuses the right food
Child has eating routine	Child has eating routine
Curative Health Investment	
Doctor has seen child at surgery	Doctor has seen child at surgery
Doctor called to home	Doctor called to home
Child had cough medicine	Child had cough medicine
Child had antibiotics	Child had antibiotics
Child had diarrhea medicine	Child had diarrhea medicine

Table 4: Measurements used for Mother Abilities and Birth Condition

Birth Condition	**Mother Socio-emotional**	**Mother Health**
Birth weight	Interpersonal awareness	Health up to pres pregnancy
Head circumference	Need for approval	History of hypertension
Crown-heel length	Separation anxiety	History of diabetes
placental weight	Timidity	Body Mass Index
	Fragile inner-self	Problems requiring regular treatment

development of the child's socio-emotional abilities. Regarding the importance of initial conditions θ_0 on the development of the child, there seems to be no discernible effect on the evolution of cognitive skills and a small effect on the behavior of the child. Omitting this factor from the analysis does not change substantially the results; for this reasons we do not take it into account for the rest of the analysis.[30]

Table 5: Technology of Cognitive and Socio-emotional Skill Formation.

		(1)	(2)	(3)	(4)
		Cognitive Skills θ^C_{t+1}		Socio-emotional Skills θ^{SE}_{t+1}	
Cognitive Factor θ^C_t	γ_1	0.622 (0.005)	0.623 (0.005)	0.000 (0.006)	0.002 (0.006)
Socio-emotional Factor θ^{SE}_t	γ_2	0.039 (0.006)	0.039 (0.006)	0.512 (0.007)	0.514 (0.006)
Parenting Investment θ^{PI}_t	δ_1	0.117 (0.009)	0.117 (0.009)	0.149 (0.010)	0.146 (0.010)
Mother Education θ^C_P	β_1	0.042 (0.003)	0.041 (0.003)	0.001 (0.003)	0.003 (0.003)
Mother Socio-emotional θ^{SE}_P	β_2	0.013 (0.004)	0.011 (0.004)	0.077 (0.005)	0.076 (0.004)
Birth Condition θ_0	β_3	0.009 (0.004)		0.017 (0.004)	

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

B. The health factor

Let's now evaluate what is the effect of taking into consideration also health as an integral component of the child's development and the mother's enduring characteristics. Table (6) shows the estimates of the following equation:

$$\begin{aligned}\theta^k_{t+1} &= \gamma^k_1\theta^C_t + \gamma^k_2\theta^{SE}_t + \gamma^k_3\theta^H_t + \delta^k_1\theta^{PI}_t \\ &+\beta^k_1\theta^C_P + \beta^k_2\theta^{SE}_P + \beta^k_3\theta^H_P + \eta^k_t\end{aligned}$$

for $t \in \{1, ..., 5\}$ and $k \in \{C, SE, H\}$.

We notice that there is a strong effect of health in the production of socio-emotional skills, and similarly socio-emotional abilities of the child are an important determinant of the child's health in the future periods. However there seems to be no significant correlation between the cognitive abilities of the child and her health status; although puzzling, this is in line with the results of Shakotko et al. (1980a,b).

[30]All of the following estimates have been carried out also including θ_0 as additional control and the difference in estimated coefficients is negligible. All tables are available from the author upon request.

Furthermore, we notice that mother health is a significant determinant of the child development, especially concerning socio-emotional skills and health.

Table 6: Technology of Cognitive, Socio-emotional and Health Formation

		(1) Cognitive Skills θ_{t+1}^{C}	(2) Socio-emotional Skills θ_{t+1}^{SE}	(3) Health θ_{t+1}^{H}
Cognitive	γ_1	0.620	0.015	0.002
Factor θ_t^C		(0.005)	(0.007)	(0.005)
Socio-emotional	γ_2	0.040	0.483	0.019
Factor θ_t^{SE}		(0.008)	(0.007)	(0.005)
Health	γ_3	0.015	0.163	0.720
Factor θ_t^H		(0.010)	(0.011)	(0.008)
Parenting	δ_1	0.125	0.139	-0.013
Investment θ_t^{PI}		(0.010)	(0.012)	(0.008)
Mother	β_1	0.043	0.003	-0.011
Education θ_P^C		(0.004)	(0.012)	(0.009)
Mother	β_2	0.011	0.065	0.016
Socio-emotional θ_P^{SE}		(0.006)	(0.006)	(0.004)
Mother	β_3	0.006	0.027	0.042
Health θ_P^H		(0.006)	(0.022)	(0.006)

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

C. Investment in Health

Parental decisions about child health care can play a fundamental part in the evolution of the child's health and wellbeing. We consider two types of health investment: preventive health care (θ_t^{PH}), and curative health care (θ_t^{CH}).[31] The first is proxied by feeding patterns, while the second is measured using the number of visits and calls to the doctor as well as the use of medicines. Using these information, we estimate the following equation:

$$\begin{aligned} \theta_{t+1}^{k} &= \gamma_1^k\theta_t^C + \gamma_2^k\theta_t^{SE} + \gamma_3^k\theta_t^H + \delta_1^k\theta_t^{PI} + \delta_2^k\theta_t^{PH} + \delta_3^k\theta_t^{CH} \\ &\quad + \beta_1^k\theta_P^{SE} + \beta_2^k\theta_P^{SE} + \beta_3^k\theta_P^H + \eta_t^k \end{aligned}$$

for $t \in \{1, ..., 5\}$ and $k \in \{C, SE, H\}$.

Table (7) shows the estimates of the effect of introducing preventive investment in the technology of skill formation. We notice that preventive health care, measured

[31] These represent the two most relevant health investments performed by the family. Sometimes, both of these types are referred to as prevention, as in Breslow (1999): "Primary prevention means averting the occurrence of a disease ...[and] ...secondary prevention means halting the progression of a disease from its early unrecognized stage to a more severe one".

as the feeding practices of the mother since birth, are not related to the cognitive development of the child, but influence the evolution of the child's health as well as the formation of her socio-emotional skills. This is in line with the results found by Motion et al. (2001), who relate the persistence of poor feeding patterns to behavioral difficulties in infancy,[32] as well as Wiles et al. (2007), who find a connection between poor eating patterns in early childhood and hyperactivity at age 7.[33]

Table 7: Technology of Cognitive, Socio-emotional and Health Formation with Preventive Health Investment

		(1) Cognitive Skills θ^C_{t+1}	(2) Socio-emotional Skills θ^{SE}_{t+1}	(3) Health θ^H_{t+1}
Cognitive	γ_1	0.618	0.014	0.001
Factor θ^C_t		(0.005)	(0.006)	(0.005)
Socio-emotional	γ_2	0.043	0.464	0.009
Factor θ^{SE}_t		(0.008)	(0.007)	(0.005)
Health	γ_3	0.014	0.147	0.710
Factor θ^H_t		(0.010)	(0.011)	(0.008)
Parenting	δ_1	0.122	0.135	-0.014
Investment θ^{PI}_t		(0.010)	(0.011)	(0.008)
Preventive	δ_2	-0.001	0.057	0.030
Care θ^{PH}_t		(0.007)	(0.008)	(0.006)
Mother	β_1	0.044	0.002	-0.012
Education θ^C_P		(0.004)	(0.008)	(0.007)
Mother	β_2	0.012	0.064	0.016
Socio-emotional θ^{SE}_P		(0.006)	(0.006)	(0.004)
Mother	β_3	0.006	0.026	0.041
Health θ^H_P		(0.006)	(0.016)	(0.005)

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

Table (8) introduces into the estimation the effect of curative health care. An important feature of our data is that it is not biased by differential access to doctors and hospitals: the health care system in the UK is public and provides equal access to mothers and children from all backgrounds and socio-economic status. Therefore we do not have to worry about issues related to health insurance and access to

[32]Contrary to our estimates, they also find significant relations between feeding difficulties and certain items of the Denver Developmental Scale. However they do not control for initial conditions and early measurement of health, cognitive or socio-emotional skills

[33]The authors use a composite measure of "junk food". The association with hyperactivity remains after controlling for socio-economic condition, a measure of IQ and other potential confounding factors; they don't find any significant association with other behavioral measures of the Strength and Difficulties Questionnaire (SDQ).

affordable health care. Our approach is similar to the joint estimation of health status and health care utilization performed by Van Vliet and Van Praag (1987), who also model health and curative care as latent variables and use data from the Netherlands in order to avoid the problem of self-selection into health insurance.[34] With the advantage of having access to a panel data, we can introduce a dynamic component into the analysis so that we can better understand the effect of health care utilization on the evolution of health.

As expected, current health status of the child and current health care utilization are negatively related.[35] However we find that curative investment has a strong and positive effect on the future health of the child, comparable to the effect that parental investment has on the cognitive development of the baby; on top of it, we also notice that the more the mother is in contact with a doctor, the better are the socio-emotional outcomes of her child.

Table 8: Technology of Cognitive, Socio-emotional and Health Formation with Curative Health Investment

		(1) Cognitive Skills θ^C_{t+1}	(2) Socio-emotional Skills θ^{SE}_{t+1}	(3) Health θ^H_{t+1}
Cognitive Factor θ^C_t	γ_1	0.621 (0.005)	0.013 (0.005)	-0.002 (0.003)
Socio-emotional Factor θ^{SE}_t	γ_2	0.043 (0.007)	0.474 (0.006)	0.014 (0.004)
Health Factor θ^H_t	γ_3	0.026 (0.007)	0.352 (0.008)	0.899 (0.005)
Parenting Investment θ^{PI}_t	δ_1	0.121 (0.009)	0.124 (0.010)	-0.019 (0.005)
Curative Care θ^{CH}_t	δ_3	0.016 (0.006)	0.196 (0.009)	0.167 (0.006)
Mother Education θ^C_P	β_1	0.045 (0.004)	0.015 (0.011)	0.000 (0.007)
Mother Socio-emotional θ^{SE}_P	β_2	0.009 (0.006)	0.062 (0.005)	0.013 (0.003)
Mother Health θ^H_P	β_3	0.007 (0.005)	0.024 (0.015)	0.035 (0.004)

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

[34] They have access to potential exogenous shifter of health care demand, such as distance from the hospital, general practitioners per thousand inhabitants, specialists per thousand. They find that only the latter are important predictors of health care utilization. Regretfully we don't have such information in our dataset; however all respondents live in the same geographic area of Avon and therefore we do not expect those aggregate variables to be important sources of variation in our setting.

[35] The covariance between θ^H_t and θ^{CH}_t is -0.17 in the first period, -0.26 in the second, -0.30 in the third, -0.33 in the fourth and -034 in the last period. See Appendix A for more details.

Finally, table (9) jointly controls for the effect of curative and preventive health investment. We notice that introducing both types of investments slightly reduces the estimates of the coefficients associated with the health factor as well as the effect of curative health care, which nevertheless remains a stronger driver of the child's health and socio-emotional ability; all the other estimates are fairly stable.

Table 9: Technology of Cognitive, Socio-emotional and Health Formation with Preventive and Curative Health Investment

		(1) Cognitive Skills θ^C_{t+1}	(2) Socio-emotional Skills θ^{SE}_{t+1}	(3) Health θ^H_{t+1}
Cognitive	γ_1	0.619	0.014	-0.002
Factor θ^C_t		(0.005)	(0.006)	(0.003)
Socio-emotional	γ_2	0.042	0.458	0.006
Factor θ^{SE}_t		(0.006)	(0.007)	(0.004)
Health	γ_3	0.032	0.334	0.892
Factor θ^H_t		(0.006)	(0.009)	(0.007)
Parenting	δ_1	0.120	0.118	-0.016
Investment θ^{PI}_t		(0.009)	(0.010)	(0.006)
Preventive	δ_2	-0.004	0.049	0.014
Care θ^{PH}_t		(0.006)	(0.007)	(0.004)
Curative	δ_3	0.021	0.191	0.165
Care θ^{CH}_t		(0.005)	(0.009)	(0.010)
Mother	β_1	0.045	0.015	0.000
Education θ^C_P		(0.003)	(0.003)	(0.002)
Mother	β_2	0.012	0.061	0.012
Socio-emotional θ^{SE}_P		(0.005)	(0.005)	(0.003)
Mother	β_3	0.005	0.024	0.035
Health θ^H_P		(0.004)	(0.004)	(0.004)

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

IV. Later Life Outcomes

In this section we present the estimates of equation (2) and we explore how the skill factors in the final period of childhood T have an effect on relevant life-time outcomes Q_j. In a follow-up survey at the age of 16, different questions were asked regarding various aspects of the teen-age life, inquiring about health, depression, anti-social behaviors and education. We used these personal achievements to anchor the technology of skill formation on concrete and relevant outcomes that are of interest both to parents and policy-makers.

Table 10: Anchoring on Later Life Outcomes

	Very Good Health	Never Suicidal	Never Shoplifted
Cognitive Factor θ_T^C	0.022 (0.050)	0.066 (0.058)	0.132 (0.045)
Socio-emotional Factor θ_T^{SE}	0.210 (0.034)	0.052 (0.042)	0.112 (0.032)
Health Factor θ_T^H	0.100 (0.040)	0.088 (0.049)	0.172 (0.037)
Mean of Q_j	*0.66*	*0.76*	*0.93*

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors. Unconditional mean of the outcome variable reported in the bottom row.

Table (10) reports the relation between age 5 cognitive abilities θ_T^C, socio-emotional skills θ_T^{SE}, and health θ_T^H and the probability of reporting very good or excellent health;[36]the probability of having never felt that "life was not worth living";[37] the probability of being in full time education;[38] the probability of not shoplifting.[39]

We can see that health is a very important determinant of depression, anti-social behavior and general health, with a magnitude that is always bigger than the effect of cognition. Not surprisingly, socio-emotional skills are the greatest determinant of depressions symptoms, and they also have a bearing on the probability of being in good health and avoiding shoplifting.

V. Conclusion

Building on an existing model of capabilities formation, we analyzed the childhood development of three important facets of human capital: cognitive abilities, socio-emotional skills, and health. A flexible model with dynamic latent factors allowed

[36]The teenager was asked to rate her general health from "excellent" (1) to "poor" (5). We constructed an indicator for reporting either excellent or very good health, which represents about 2/3 of the population.

[37]About 76% of the 16-years-old state they never had such symptom of depression

[38]82% of the adolescents report attending school full time

[39]93% of the interviewed report has never having "taken something from a shop without paying for it in the past year". This question was asked in a previous questionnaire, at the age 14.

tackling the pervasive issue of imperfect proxies, so that reliable estimates of the interaction between these different traits could be evaluated. A linear technology of health and skill formation was estimated, taking into consideration parental investment, curative health care, preventive health care, and maternal characteristics.

Our analysis gives strong empirical support to the claim that health is a fundamental part of human capital that dynamically interacts and evolves with other skills and capabilities. An illness can slow down the socio-emotional and cognitive growth of the child, while good physical health promotes the flourishing of human capabilities. Furthermore, these processes start interacting very early in life and build on each other since birth.

One might expect that these dynamic complementarities apply to severe chronic conditions and disabilities, which certainly incapacitate the normal cognitive, social and emotional development of a child. However we demonstrate that similar relations hold for common ailments and acute, short-term illnesses that virtually any child experiences while growing up.

We also show how simple investments in curative and preventive health care that start since birth have a strong positive impact on children's health, cognitive, and especially socio-emotional skills.

Ignoring the health dimension in the process of human capital formation biases our perception and fails to properly take into account the synergies and spillovers that characterize the biological, social, and cognitive evolution of children. An improvement in any facet of human capital propagates through the complex architecture of our bodies and minds. Similarly, interventions that induce better parenting practices and greater investment can influence children in multiple ways, and impact their development in many dimensions.

Every individual is a single, interconnected organism that unifies various skills. This form of bundling represents both a problem - an educator cannot teach math to feverish children - but also a potential opportunity, since a comprehensive approach to child development can fully capitalize on the multiple synergies and interconnections among different capabilities, promoting more efficiently the wellbeing of young individuals.

However current policies for early childhood development seem to negate and neutralize such important links. Halfon et al. (2009) show how services targeting children in the U.S., England, Canada and Australia have been excessively fragmented: pediatricians, educators and psychologists work in separate environments and compartmentalized structures that focus on their area of expertize, and seldom interact with each other even when dealing with the same child.[40,41] Furthermore, the early periods of life, before the entry in kindergarten, are often neglected by national public policies, so that the burden of child care falls mainly on the family.[42]

Building on policies suggested by studies in public health,[43] our results support

[40]Interestingly enough, this was not the case a century ago. Both Margaret McMillan in England, and Maria Montessori and the Agazzi sisters in Italy, placed a great emphasis on health prevention, routine dental and physical screenings, and promotion of proper hygiene in their nursery schools. See Lascarides and Hinitz (2000).

[41]This is true also for professors and academics and not only for practitioners. For an argument in favor of greater integration of various disciplines, see Duncan (2012).

[42]Glascoe (2000) estimates that about 70% of children with developmental problems at kindergarten entry could have been identified earlier, but were not.

[43]See for example the public health strategies proposed in Halfon et al. (2003); Halfon and Inkelas (2003) and Shonkoff et al. (2009).

the notion that preventive and curative health care for children should start very early and be fully integrated into the realm of family policies and early education; screening and prevention should start since birth; and the wellbeing of the child as a whole should be the focus of policy, not simply the promotion or prevention of a very specific outcomes measures such as reading ability or obesity. Only this holistic approach can fully capture the biological and technological synergies between health, socio-emotional, and cognitive development of the child.

Acknowledgments

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and Pietro Biroli will serve as guarantors for the contents of this paper.

References

Ackerman, P. L. and E. D. Heggestad (1997): "Intelligence, personality, and interests: evidence for overlapping traits." *Psychological bulletin*, 121, 219–45.

Almlund, M., A. L. Duckworth, J. J. Heckman, and T. Kautz (2011): "Personality Psychology and Economics," in *Handbook of the Economics of Education*, ed. by E. Hanushek, Amsterdam, The Netherlands: North Holland, vol. 4, chap. 1, 1–181, 4 ed.

Anderson, T. W. and H. Rubin (1956): "Statistical inference in factor analysis," in *Proceedings of the third Berkeley Symposium on mathematical statistics and probability: contributions to the theory of statistics*, University of California Press, vol. V, 111.

Barker, D. J. P., C. Osmond, J. Golding, D. Kuh, and M. E. J. Wadsworth (1989a): "Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease disease," *British Medical Journal*, 298, 564–567.

Barker, D. J. P., C. Osmond, P. D. Winter, B. Margetts, and S. J. Simmonds (1989b): "Weight in infancy and death from ischaemic heart disease," *The Lancet*, 8663, 577–580.

Belsky, J., B. Bell, R. H. Bradley, N. Stallard, and S. Stewart-Brown (2007): "Socioeconomic risk, parenting during the preschool years and child health age 6 years." *European Journal of public health*, 17, 508–13.

Bleakley, H. (2010): "Health, Human Capital, and Development," *Annual Review of Economics*, 2, 283–310.

Bollen, K. A. (1989): *Structural Equations with Latent Variables*, New York, NY: Wiley Series in Probability and Mathematical Statistics.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. Ter Weel (2008): "The economics and psychology of personality traits," *The Journal of Human Resources*, XLIII.

Bowles, S. and H. Gintis (2002): "The inheritance of inequality," *The Journal of Economic Perspectives*, 16, 3–30.

Boyce, P. and G. Parker (1989): "Development of a scale to measure interpersonal sensitivity." *The Australian and New Zealand journal of psychiatry*, 23, 341–51.

Boyd, A., J. Golding, J. Macleod, D. A. Lawlor, A. Fraser, J. Henderson, L. Molloy, A. R. Ness, S. M. Ring, and G. Davey Smith (2013): "Cohort Profile: The 'Children of the 90s'–the index offspring of the Avon Longitudinal Study of Parents and Children." *International Journal of Epidemiology*, 42, 111–127.

Breslow, L. (1999): "From disease prevention to health promotion," *JAMA: the Journal of the American Medical Association*, 281, 1030–1033.

Brown, T. A. (2006): *Confirmatory Factor Analysis for Applied Research*, New York, NY: The Guilford Press.

Caldwell, B. M. and R. H. Bradley (1984): "Home Observation for Measurement of the Environment: Administration Manual," *Family & Human Dynamics Research Institute, Arizona State University.*

Carey, W. B. and S. C. McDevitt (1977): "Infat Temperament Questionnaire (4-8 months)," *Dept. Educational Psychology, Temple University.*

Carneiro, P., K. T. Hansen, and J. J. Heckman (2003): "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice," *International Economic Review*, 44, 361–422.

Carneiro, P., C. Meghir, and M. Parey (2013): "Maternal Education , Home Environments and the Development of Children and Adolescents," *Journal of the European Economic Association*, 11, 123–160.

Case, A., A. Fertig, and C. Paxson (2005): "The lasting impact of childhood health and circumstance." *Journal of Health Economics*, 24, 365–89.

Case, A., D. Lubotsky, and C. Paxson (2002): "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 92, 1308–1334.

Case, A. and C. Paxson (2002): "Parental Behavior And Child Health," *Health Affairs*, 21, 164–178.

Cattell, R. B. (1966): "The Scree Test For The Number Of Factors," *Multivariate Behavioral Research*, 1, 245–276.

Cawley, J., J. J. Heckman, and E. Vytlacil (2001): "Three observations on wages and measured cognitive ability," *Labour Economics*, 8, 419–442.

Chatterji, P., K. Lahiri, and J. Song (2012): "The Dynamics Of Income-Related Health Inequality Among American Children," *Health economics*.

Condliffe, S. and C. R. Link (2008): "The relationship between economic status and child health: Evidence from the United States," *American Economic Review*, 98, 1605–1618.

Conti, G. and J. J. Heckman (2010): "Understanding the Early Origins of the Education-Health Gradient: A Framework That Can Also Be Applied to Analyze Gene-Environment Interactions," *Perspectives on Psychological Science*, 5, 585–605.

Conti, G., J. J. Heckman, and S. Urzua (2010): "The Education-Health Gradient," *American Economic Review: Papers & Proceedings*, 100, 234–38.

Costello, A. B. and J. W. Osborne (2005): "Best Practices in Exploratory Factor Analysis : Four Recommendations for Getting the Most From Your Analysis," *Practical Assessment, Research & Evaluation*, 10.

Cronbach, L. J. (1951): "Coefficient alpha and the internal structure of tests," *Psychometrika*, 16, 297–334.

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

——— (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *The Journal of Human Resources*, 43, 738–782.

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

Cunha, F., J. J. Heckman, and S. Schennach (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica*, 78, 883–931.

Currie, A., M. A. Shields, and S. W. Price (2007): "The child health/family income gradient: Evidence from England." *Journal of Health Economics*, 26, 213–32.

Currie, J. (2009): "Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development," *Journal of Economic Literature*, 112, 3–56.

Currie, J. and D. Almond (2011): "Human capital development before age five," in *Handbook of Labor Economics*, National Bureau of Economic Research Cambridge, Mass., USA, November 2009, 1315–1486.

Currie, J. and E. Moretti (2007): "Biology as Destiny ? Short- and Long- Run Determinants of Intergenerational Transmission of Birth Weight," *Journal of Labor Economics*, 25.

Currie, J. and M. Stabile (2003): "Socioeconomic status and child health: Why is the relationship stronger for older children?" *American Economic Review*, 93, 1813–1823.

Currie, J., M. Stabile, P. Manivong, and L. L. Roos (2010): "Child Health and Young Adult Outcomes," *Journal of Human Resources*, 45, 517–548.

Danese, A., A. Caspi, B. S. Williams, A. Ambler, K. Sugden, J. T. Mika, H. Werts, E. J. Freeman, C. M. Pariante, T. E. Moffitt, and L. Arseneault (2011): "Biological embedding of stress through inflammation processes in childhood." *Molecular psychiatry*, 16, 244–6.

Danese, A., T. E. Moffitt, H. Harrington, B. J. Milne, G. Polanczyk, C. M. Pariante, R. Poulton, and A. Caspi (2009): "Adverse childhood experiences and adult risk factors for age-related disease: depression, inflammation, and clustering of metabolic risk markers." *Archives of pediatrics & adolescent medicine*, 163, 1135–43.

Duggan, A., E. C. McFarlane, A. M. Windham, C. A. Rohde, D. S. Salkever, L. Fuddy, L. A. Rosenberg, S. B. Buchbinder, and C. C. Sia (1999): "Evaluation of Hawaii's healthy start program," *The Future of Children*, 9, 66–90.

Duncan, G. J. (2012): "Give us this day our daily breadth." *Child development*, 83, 6–15.

Duncan, G. J. and K. A. Magnuson (2005): "Can family socioeconomic resources account for racial and ethnic test score gaps?" *The Future of Children*, 15, 35–54.

Elander, J. and M. Rutter (1996): "Use and development of the Rutter parents' and teachers' scales." *Journal of Methods in Psychiatric Research*, 6, 63–78.

Fabrigar, L. R. (2011): *Exploratory Factor Analysis*, Oxford University Press.

Fenson, L., P. S. Dale, and S. Reznic (1991): "The MacArthur Communicative Development Inventories: User's Guide and Technical Manual," *Developmental Psuchology Laboratory, San Diego, CA.*

Frankenburg, W. K. and J. B. Dodds (1967): "The Denver developmental screening test." *The Journal of pediatrics*, 71, 181–91.

Fraser, A., C. Macdonald-Wallis, K. Tilling, A. Boyd, J. Golding, G. Davey Smith, J. Henderson, J. Macleod, L. Molloy, A. R. Ness, S. M. Ring, S. M. Nelson, and D. A. Lawlor (2013): "Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort." *International Journal of Epidemiology*, 42, 97–110.

Fullard, W., S. C. McDevitt, and W. B. Carey (1978): "Toddler Temperament Scale," *Dept. Educational Psychology, Temple University.*

Glascoe, F. P. (2000): "Evidence-based approach to developmental and behavioural surveillance using parents' concerns," *Child: Care, Health and Development*, 26, 137–149.

Gluckman, P. and M. Hanson (2006): "The Developmental Origins of Health and Disease," in *Early Life Origins of Health and Disease*, Springer, 1–7.

Goodman, A., R. Joyce, and J. P. Smith (2011): "The long shadow cast by childhood physical and mental problems on adult life." *Proceedings of the National Academy of Sciences*, 108, 6032–7.

Grossman, M. (1972): "On the concept of health capital and the demand for health," *Journal of Political Economy*, 80, 223–255.

Halfon, N. and M. Inkelas (2003): "Optimizing the health and development of children." *Journal of the American Medical Association*, 290, 3136–8.

Halfon, N., M. Regalado, K. T. McLearn, A. A. Kuo, and K. Wright (2003): "Building a bridge from birth to school: Improving developmental and behavioral health services for young children," Tech. Rep. 564, The Commonwealth Fund.

Halfon, N., S. Russ, F. Oberklaid, J. Bertrand, and N. Eisenstadt (2009): "An international comparison of early childhood initiatives: From services to systems," *Journal of Developmental & Behavioral Pediatrics*, 30.

Hampson, S. E. (2012): "Personality processes: mechanisms by which personality traits "get outside the skin"." *Annual review of Psychology*, 63, 315–39.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004): "The effect of schooling and ability on achievement test scores," *Journal of Econometrics*, 121, 39–98.

Heath, A. C., K. Berg, L. J. Eaves, M. H. Solaas, L. A. Corey, J. K. Sundet, P. Magnus, and W. E. Nance (1985): "Education policy and the heritability of educational attainment," *Nature*, 314, 734–736.

Heckman, J. J. (2007): "The economics, technology, and neuroscience of human capability formation." *Proceedings of the National Academy of Sciences*, 104, 13250–5.

——— (2012): "The developmental origins of health," *Health Economics*, 29, 24–29.

Heckman, J. J. and T. Kautz (2012): "Hard evidence on soft skills," *Labour Economics*, 19, 451–464.

HECKMAN, J. J., S. H. MOON, R. PINTO, P. A. SAVELYEV, AND A. YAVITZ (2010): "Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program," *Quantitative Economics*, 1, 1–46.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24, 411–482.

HERTZMAN, C. AND T. W. BOYCE (2010): "How experience gets under the skin to create gradients in developmental health." *Annual Review of Public Health*, 31, 329–347.

JOKELA, M., G. D. BATTY, I. J. DEARY, C. R. GALE, AND M. KIVIMÄKI (2009): "Low childhood IQ and early adult mortality: the role of explanatory factors in the 1958 British Birth Cohort." *Pediatrics*, 124, e380–8.

JOLIFFE, I. T. AND B. MORGAN (1992): "Principal component analysis and exploratory factor analysis," *Statistical Methods in Medical Research*, 1, 69–95.

KAUTZ, T., J. J. HECKMAN, R. DIRIS, B. TER WEEL, AND L. BORGHANS (2014): "Fostering and Measuring Skills: improving cognitive and non-cognitive skills to promote lifetime success," *OECD*.

LASCARIDES, C. V. AND B. F. HINITZ (2000): *History of Early Childhood Education*, New York: Falmer Press.

LEVITT, P. (2003): "Structural and functional maturation of the developing primate brain." *The Journal of pediatrics*, 143, S35–45.

LOPES, H. F. AND S. FRUHWIRTH-SCHNATTER (2010): "Parsimonious Bayesian Factor Analysis when the Number of Factors is Unknown," *Working Paper*, 1–37.

MARMOT, M. (2010): "Fair society, healthy lives: the Marmot review; strategic review of health inequalities in England post-2010," .

MARSHALL, A. (1890): *Principles of Economics*, Macmillan and Co., v. 1 ed.

MERRICK, J. AND E. CARMELI (2003): "A review on the prevalence of disabilities in children," *The Internet Journal of Pediatrics and Neonatology*, 3.

MOTION, S., K. NORTHSTONE, AND A. EMOND (2001): "Persistent early feeding difficulties and subsequent growth and developmental outcomes," *Ambulatory Child Health*, 231–237.

MUENNIG, P., L. J. SCHWEINHART, J. MONTIE, AND M. NEIDELL (2009): "Effects of a prekindergarten educational intervention on adult health: 37-year follow-up results of a randomized controlled trial." *American Journal of Public Health*, 99, 1431–7.

NUSSBAUM, M. C. (2011): *Creating Capabilities: The Human Development Approach*, Cambridge, MA: Harvard University Press.

OLDS, D. L. (2002): "Prenatal and infancy home visiting by nurses: from randomized trials to community replication." *Prevention Science*, 3, 153–72.

PALLONI, A., C. MILESI, R. G. WHITE, AND A. TURNER (2009): "Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials." *Social Science & Medicine*, 68, 1574–82.

SCHENNACH, S. (2004): "Estimation of Nonlinear Models with Measurement Error," *Econometrica*, 59, 1161–75.

SEN, A. K. (1985): "Well-being, agency and freedom: the Dewey lectures 1984," *The Journal of Philosophy*, 82, 169–221.

——— (1990): "Development as capability expansion," in *Human development and the international development strategy for the 1990s*, ed. by K. B. Griffin and J. E. Knight, London: Macmillan Publishers, chap. 2, 41–58.

SHAKOTKO, R. A., L. N. EDWARDS, AND M. GROSSMAN (1980a): "An exploration of the dynamic relationship between health and cognitive development in adolescence," *NBER working paper*.

SHAKOTKO, R. A., L. LILLARD, AND R. MOFFITT (1980b): "Dynamic Aspect of Children's Health, Intellectual Development, and Family Economic Status," *NBER working paper*.

Shonkoff, J. P., T. W. Boyce, and B. S. McEwen (2009): "Neuroscience, molecular biology, and the childhood roots of health disparities," *JAMA: the Journal of the American Medical Association*, 301.

Skrondal, A. and P. Laake (2001): "Regression among factor scores," *Psychometrika*, 66, 563–575.

Smith, J. P. (2009): "The Impact of Childhood Health on Adult Labor Market Outcomes," *Review of Economics and Statistics*, 91, 478–489.

Stewart-Brown, S., L. Fletcher, and M. E. J. Wadsworth (2005): "Parent-child relationships and health problems in adulthood in three UK national birth cohort studies." *European Journal of public health*, 15, 640–6.

Van Vliet, R. C. and B. Van Praag (1987): "Health status estimation on the basis of MIMIC-health care models," *Journal of Health Economics*, 6, 27–42.

Waylen, A., N. Stallard, and S. Stewart-Brown (2008): "Parenting and health in mid-childhood: a longitudinal study." *European Journal of public health*, 18, 300–5.

Wiles, N. J., K. Northstone, and P. M. Emmett (2007): "'Junk food' diet and childhood behavioural problems: results from the ALSPAC cohort," *European Journal of clinical nutrition*, 63, 491–498.

Williams, B. D. (2013): "A Measurement Model with Discrete Measurements and Continuous Latent Variables," *Working Paper*.

A. Latent Factor Distribution

In this Appendix there is a more detailed description of the distribution of the latent factors estimated using the ALSPAC sample, as described in section II.B. Here below are the estimated Variance-Covariance matrix of all the factors for each time period.

Table A1: Variance-Covariance Matrix of the factors, time 1

	θ_1^C	θ_1^{SE}	θ_1^H	θ_1^{PI}	θ_1^{PH}	θ_1^{CH}	θ_P^{SE}	θ_P^H	θ_P^C
Cognitive	**0.349**	0.064	-0.003	0.058	0.015	0.031	0.028	-0.001	-0.126
Socio-emotional		**0.460**	0.074	0.065	0.140	-0.058	0.077	0.059	-0.009
Health			**0.189**	0.033	0.100	-0.173	0.055	0.077	0.006
Parenting				**0.192**	0.049	0.003	0.034	0.049	0.021
Preventive Care					**0.216**	-0.092	0.061	0.063	0.093
Curative Care						**0.218**	-0.040	-0.066	-0.057
Mom Socio-emotional							**0.767**	0.108	-0.040
Mom Health								**0.984**	0.078
Mom Education									**1.585**

Table A2: Variance-Covariance Matrix of the factors, time 2

	θ_2^C	θ_2^{SE}	θ_2^H	θ_2^{PI}	θ_2^{PH}	θ_2^{CH}	θ_P^{SE}	θ_P^H	θ_P^C
Cognitive	**0.537**	0.095	0.004	0.113	0.026	0.017	0.034	0.019	-0.006
Socio-emotional		**0.558**	0.095	0.085	0.135	-0.064	0.102	0.080	0.015
Health			**0.278**	0.022	0.078	-0.260	0.056	0.094	0.000
Parenting				**0.257**	0.029	0.007	0.043	0.047	0.050
Preventive Care					**0.517**	-0.061	0.083	0.059	0.008
Curative Care						**0.340**	-0.034	-0.075	-0.078
Mom Socio-emotional							**0.767**	0.108	-0.040
Mom Health								**0.984**	0.078
Mom Education									**1.585**

Table A3: Variance-Covariance Matrix of the factors, time 3

	θ_3^C	θ_3^{SE}	θ_3^H	θ_3^{PI}	θ_3^{PH}	θ_3^{CH}	θ_P^{SE}	θ_P^H	θ_P^C
Cognitive	**0.622**	0.116	0.010	0.150	0.024	0.008	0.039	0.032	0.073
Socio-emotional		**0.598**	0.141	0.096	0.144	-0.069	0.118	0.094	0.021
Health			**0.352**	0.013	0.075	-0.302	0.058	0.108	-0.010
Parenting				**0.301**	0.019	0.009	0.050	0.046	0.074
Preventive Care					**0.632**	-0.046	0.096	0.056	-0.044
Curative Care						**0.404**	-0.030	-0.082	-0.092
Mom Socio-emotional							**0.767**	0.108	-0.040
Mom Health								**0.984**	0.078
Mom Education									**1.585**

B. Construction of the Factors

B1. The Measurements

A fundamental part of the analysis is the choice of the measures that construct the latent factors analyzed throughout the paper. Although most of the analysis

Table A4: Variance-Covariance Matrix of the factors, time 4

	θ_4^C	θ_4^{SE}	θ_4^H	θ_4^{PI}	θ_4^{PH}	θ_4^{CH}	θ_P^{SE}	θ_P^H	θ_P^C
Cognitive	**0.665**	0.131	0.016	0.175	0.020	0.002	0.044	0.041	0.124
Socio-emotional		**0.634**	0.193	0.101	0.152	-0.072	0.127	0.103	0.018
Health			**0.418**	0.004	0.077	-0.321	0.060	0.119	-0.023
Parenting				**0.329**	0.013	0.009	0.055	0.045	0.093
Preventive Care					**0.677**	-0.038	0.104	0.055	-0.075
Curative Care						**0.437**	-0.027	-0.086	-0.103
Mom Socio-emotional							**0.767**	0.108	-0.040
Mom Health								**0.984**	0.078
Mom Education									**1.585**

Table A5: Variance-Covariance Matrix of the factors, time 5

	θ_5^C	θ_5^{SE}	θ_5^H	θ_5^{PI}	θ_5^{PH}	θ_5^{CH}	θ_P^{SE}	θ_P^H	θ_P^C
Cognitive	**0.688**	0.142	0.023	0.192	0.017	-0.002	0.048	0.047	0.158
Socio-emotional		**0.670**	0.242	0.104	0.157	-0.073	0.135	0.110	0.012
Health			**0.478**	-0.003	0.081	-0.329	0.063	0.128	-0.037
Parenting				**0.349**	0.009	0.009	0.059	0.044	0.108
Preventive Care					**0.695**	-0.034	0.109	0.054	-0.094
Curative Care						**0.455**	-0.025	-0.090	-0.111
Mom Socio-emotional							**0.767**	0.108	-0.040
Mom Health								**0.984**	0.078
Mom Education									**1.585**

presented in the paper is carried out at the level of the unobservable factors, the observable measures that lie underneath are the raw source of data and variation that allow us to obtain the main results. Therefore selection of the measures to use has been done with extreme care.

We initially selected all the potentially relevant variables from the mother's and child's questionnaires, which were already subdivided into sections that are related to our different factors.[44] When choosing which measurements to use, we followed both a theoretical and a statistical approach. A priori, we privileged the inclusion of variables that had been validated and widely used in the psychometric and psychological literature, as well as questions that were asked consistently over time. This was done in order to be comparable with the existing literature and to incorporate all the dynamic feature of the raw data. In the analysis of the data, we used Exploratory Factor Analysis (EFA)[45] in order to evaluate the stability of each latent factor and to discard those measures that were poorly related to the construct of interest. We then performed Confirmatory Factor Analysis (CFA)[46] and included into the general model all of the variables previously selected, estimating jointly both the measurement system (3) and the dynamic technology of skill formation (1). Following the suggestions of Costello and Osborne (2005) and for parsimony of the estimation[47], we discarded from the measurement system those variables that dis-

[44] For example, some names of the sections are "About your baby: milestones" and "Understanding and Talking"; "Temperament"; "Your Baby's Health"; "You and your baby" and "Looking after your baby" and "Your infant and her environment"; "Feeding"; "Problems and treatment: doctors". They contained respectively the measures that we used for constructing the cognitive; socio-emotional; health; investment; preventive health; and curative health factors.

[45] For a discussion of its use, see among others Fabrigar (2011); Joliffe and Morgan (1992)

[46] See Bollen (1989); Brown (2006)

[47] The number of parameters estimated to produce Table 9 are already 1,011

played either a very low factor loading, indicating a poor correlation with the other measures and the latent factor (the *weak-loading problem*), or a low communality, indicating a poor noise-to-signal ratio[48].

B2. Number of factors

The dimensionality of the stock of human capital is a highly debated issue among professionals of different disciplines. Considering cognition in adulthood, psychometricians have developed a hierarchy of "orders" among different mental functions: a high order factor (sometimes called general factor g) is predictive of all cognitive tasks, while many lower order factors are predictive of performance in particular tasks, like verbal ability, numeracy, coding speed, etc.[49]. Considering adult personality, a partial consensus within personality psychologists was reached in the construction of a 5-factor model (The Big Five), but other competing models with higher dimensionality are still used.[50] Furthermore, one could theorize that the number of factors increases with age: a small number of factors could characterize fairly well the facets of human capital of a toddler, but a greater variety of cognitive, health, and personality traits would better capture the human capital of an adult. New traits could be flourishing over the life-cycle.

Since we are considering only the very first years of life, and the purpose of the paper is to characterize broad linkages across different domains of human capital, we consider a parsimonious and tractable model with only three main factors. Such theoretical decision is also backed-out by the measurement present in the data: figure (B1) is a screeplot of the eigenvalues of the covariance matrix of all the measurement used in each time period. Following the suggestions from Cattell (1966), the optimal number of factors should be chosen when an elbow appears on the curve. Although not always strikcly evident, the choice of 6 factors (Cognitive, Socio-emotional, Health, Parenting, Curative and Preventive care) seems reasonable also from an empirical perspective.

B3. Reliability and Consistency

A first approximation of the reliability and the internal consistency of the measures chosen can be summarized by Cronbach's α (Cronbach (1951)), a very widely used statistic calculated from the pairwise correlations between items.

As shown in table ($B1$), most of the factors used in the paper have a quite high internal consistency. A lower level of reliability is found among the measurements used for the first period, from 0 to 4 months of life, and in the construction of the curative health care factor.

A more in depth analysis requires moving away from the latent constructs and analyzing directly the raw measures that are at the heart of the factor analysis.

In order to have a better idea of the importance of each measure in the construction of the latent factors, we report some of the relevant estimates that are essential for the identification of the model used. Consider the general measurement equation for factor k at period t

$$M_t^{k,l} = \phi_t^{k,l} X_t + \alpha_t^{k,l} \theta_t^k + \varepsilon_t^{k,l}$$

[48]See equation (B1) below for a more precise definition of communality. Tables ($B2 - B9$) report the estimated factor loadings and commonalities of the measures included

[49]See for instance Ackerman and Heggestad (1997)

[50]See Almlund et al. (2011) for a discussion.

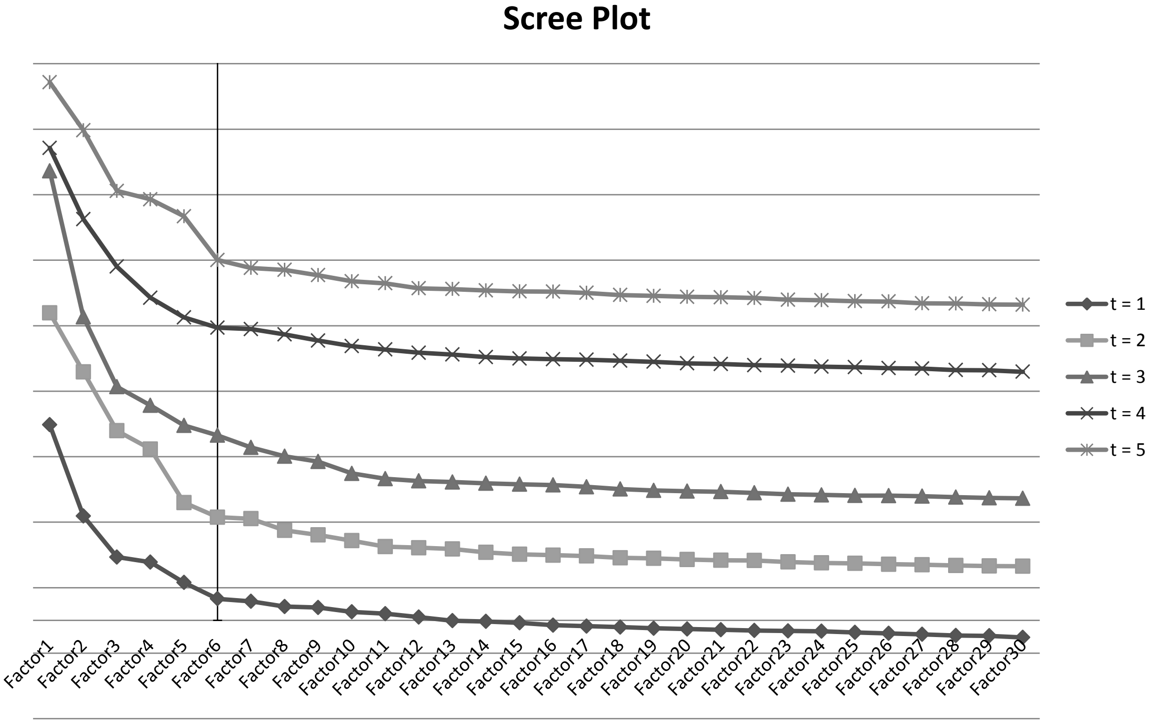

Figure B1: Scree Plot

netting out the variation that is due to the observable covariates X_t, the variance of each measurement can be decomposed into the part that is related to the latent factor (signal) and the part that is due to the error term (noise):

$$Var\left(M_t^{k,l}|X_t\right) = \underbrace{\left(\alpha_t^{k,l}\right)^2 Var(\theta_t^k)}_{\text{signal}} + \underbrace{Var(\varepsilon_t^{k,l})}_{\text{noise}}$$

With this distinction in mind we can compute the communality of each measure, which represents the share of the variance that can be attributed to the signal:

Table B1: Internal Consistency of the Factors: Cronbach's α

	t = 1	t = 2	t = 3	t = 4	t = 5
Cognitive Factor	0.641	0.790	0.859	0.800	0.812
Socio-emotional Factor	0.801	0.795	0.688	0.551	0.666
Health Factor	0.442	0.607	0.623	0.596	0.607
Parental Investment	0.569	0.558	0.649	0.551	0.662
Preventive Health Investment	0.350	0.829	0.667	0.658	0.779
Curative Health Investment	0.171	0.489	0.457	0.404	0.395
Mother Socio-emotional	0.837	Mother Health	0.362	Initial Condition θ_0	0.840

$$s_{\theta}^{t,k,l} = \frac{\left(\alpha_t^{k,l}\right)^2 Var(\theta_t^k)}{\left(\alpha_t^{k,l}\right)^2 Var(\theta_t^k) + Var(\varepsilon_t^{k,l})} \tag{B1}$$

The uniqueness, which is the share of the variance that can be attributed to the noise, is simply the residual share of the variance: $s_{\varepsilon}^{t,k,l} = 1 - s_{\theta}^{t,k,l}$. The higher the communality, the more the measure is relevant for the construction of the factor and the lower is the measurement error intrinsic in the proxy used.

Tables $(B2 - B9)$ report the estimates of the factor loadings $\alpha_t^{k,l}$ as well as the commonalities $s_{\theta}^{t,k,l}$ for all the measures used in the paper. As we can see, the prevalence of measurement error varies a lot over different factors k and over time t; however we cut the data, we see that the noise is quite substantial, especially in earlier measures as well as in those variables that are associated to the investment. The fact that a high share of the overall is due to noise further substantiates the need for an appropriate model that takes this feature of the data into account: the dynamic factor model used in this paper is our suggestion to tackle this problem.

Table B2: Cognitive Factors: loadings and signal

Measurement	α	s_θ	**Measurement**	α	s_θ
t = 1			**t = 3**		
Baby laughs	1.000	35.95%	Vocabulary knowledge	1.000	77.91%
Baby looks at mum's face	0.583	11.69%	Language knowledge	0.896	56.61%
Baby follows mum with eyes	0.806	22.62%	Plurals knowledge	0.942	62.25%
Baby smiles	1.050	38.45%	Grammar	0.870	54.37%
Baby squeals	0.809	23.14%	Communication	0.813	47.12%
Baby lifts head when on tummy	0.489	8.32%	Social achievement	0.554	20.77%
Baby touches hands together	0.529	10.03%	Social achievement	0.464	15.65%
			Fine motor	0.435	12.49%
t = 2			**t = 4**		
Vocabulary knowledge	1.000	56.57%	Vocabulary knowledge	1.000	64.70%
Understanding	0.855	41.23%	Past tense knowledge	0.942	57.62%
Non-verbal communication	0.768	34.01%	Plurals knowledge	0.804	41.89%
Communication	0.553	16.69%	Word combination	0.958	63.59%
Social skills	0.663	24.12%	Social development	0.560	23.21%
Social development	0.900	45.82%	Fine motor	0.607	27.09%
Fine motor	0.673	25.71%	**t=5**		
Gross motor	0.535	15.81%	Reading and counting	1.000	55.84%
			Playing and sharing	0.681	33.52%
			Listening and singing	0.806	43.88%
			Social skills	0.851	51.73%
			Drawing skills	0.836	51.96%
			Building skills	0.522	23.99%

B4. Dedicated Measurements

Throughout the paper we maintain the assumption of dedicated measurements, which means we assume that each measure is related to one factor only and does not load on any other trait. Considering the measurement equation $\mathbf{M}_t = \mathbf{\Phi}_t\mathbf{X}_t + \mathbf{\Lambda}_t\boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t$, this is equivalent to imposing restrictions on the matrix of factor loadings $\mathbf{\Lambda}$, notably that each row has only one non-zero entry. For example, considering measure

Table B3: Socio-emotional Factors: loadings and signal

Measurement	α	s_θ	Measurement	α	s_θ
t = 1			**t = 3**		
Baby is grizzly	1.000	48.81%	Adaptability	1.000	60.98%
Baby is placid	0.930	40.09%	Approach	0.346	7.24%
Baby is fretful	1.049	50.40%	Mood	0.972	56.28%
Baby is demanding	0.889	37.13%	Rhythmicity	0.404	9.75%
Baby is angry	0.903	37.86%	Persistence	0.557	18.76%
Baby is stubborn	0.637	18.87%	Activity	0.651	25.46%
Baby is happy	0.785	28.27%	Intensity	0.655	25.57%
			t = 4		
t = 2			Conduct difficulties	1.000	57.15%
Adaptability	1.000	55.75%	Emotional difficulties	0.548	19.03%
Approach	1.004	55.93%	Hyperactivity	0.719	32.71%
Mood	0.983	53.08%	Prosocial	0.452	13.27%
Rhythmicity	0.566	17.88%	**t=5**		
Persistence	0.625	21.76%	Conduct difficulties	1.000	66.59%
Distractibility	0.929	47.53%	Activity	0.542	19.69%
			Aggressivity	0.659	30.35%
			Relation with other children	0.682	31.37%
			Sociability	0.501	17.01%
			Concentration	0.344	8.02%

Table B4: Health Factors: loadings and signal

Measurement	α	s_θ	Measurement	α	s_θ
t = 1			**t = 4**		
Health of child in past month	1.000	19.88%	Health of child in past month	1.000	40.81%
Baby vomits	0.941	16.86%	Health of child in past year	0.956	36.02%
Baby possets	0.657	8.22%	Child had .. cough	0.461	8.79%
Baby ever had .. cough	0.437	3.61%	... vomiting	0.422	7.34%
... high temperature	0.559	5.94%	... stomach ache	0.527	11.46%
... sniffles	0.836	13.27%	... high temperature	0.731	21.65%
t = 2			... earache	0.623	15.81%
Health of child in past month	1.000	28.47%	... rash	0.415	7.12%
Health of child in 1st month	1.082	32.64%	... headache	0.405	6.78%
Child had .. cough	0.783	17.16%			
... diarrhea	0.677	12.72%	**t=5**		
... vomiting	0.802	17.82%	Health of child in past month	1.000	43.54%
... high temperature	0.795	17.76%	Health of child in past year	0.819	30.09%
... cold	0.497	6.89%	Child had .. cough	0.410	7.92%
t = 3			... vomiting	0.480	10.80%
Health of child in past month	1.000	52.56%	... stomach ache	0.509	12.13%
Child had .. cough	0.539	20.51%	... high temperature	0.675	20.84%
... diarrhea	0.513	14.20%	... earache	0.686	21.50%
... vomiting	0.594	19.36%	... rash	0.445	9.29%
... stomach ache	0.507	14.54%	... headache	0.470	10.31%
... high temperature	0.671	22.07%	Number of infections	0.813	29.91%
... earache	0.454	14.96%			

Table B5: Parental Investment: loadings and signal

Measurement	α	s_θ	Measurement	α	s_θ
t = 1			**t = 4**		
Feelings about pregnancy	1.000	21.72%	Mum plays games with child	1.000	29.22%
Reaction to pregnancy	1.156	26.32%	Taken to library	0.460	6.97%
Motherhood means personal sacrifice	0.615	7.56%	Taken to interesting places	0.504	8.46%
Motherhood gives new opportunities	0.956	18.98%	Mum talks to child while working	0.744	18.09%
t = 2			Mum reads to the child	0.925	28.28%
Mum shows child picture books	1.000	26.65%	Mom teaches child	0.769	19.88%
Mum plays games with child	0.736	14.03%	Maternal Enjoyment	0.724	17.35%
Mum and child play with toys	0.769	15.55%	Maternal Confidence	0.555	10.20%
Mum talks to child while working	0.885	20.17%			
No of books child owns	0.733	14.33%	**t = 5**		
Mum tries to teach child	0.546	7.79%	Mum makes things with child	1.000	37.84%
Maternal bonding	0.715	13.25%	Mum plays with toys with child	1.121	44.26%
t = 3			Mum draws or paints with child	1.092	42.23%
Mum plays games with child	1.000	27.78%	Taken to library	0.401	5.60%
Mum plays with toys with child	0.902	24.82%	Taken to interesting places	0.425	6.37%
Taken to interesting places	0.542	10.95%	Mum reads to the child	0.805	22.77%
Taken to library	0.413	6.20%	Mum talks to child while working	0.715	17.82%
Mum talks to child while working	0.669	17.95%			
Mum reads to the child	1.007	40.44%			
Mom teaches child	0.860	31.46%			
No of stimulating toys	0.715	16.15%			
Maternal experience	0.628	11.88%			

Table B6: Preventive Health Investment: loadings and signal

Measurement	α	s_θ	Measurement	α	s_θ
t = 1			**t = 3**		
Difficulties feeding	1.000	21.64%	Child been choosy with food	1.000	63.21%
Breastfeeding duration	0.429	3.98%	Child refuses the right food	0.970	59.45%
Child refuses milk	0.229	1.13%	Child has eating routine	0.559	19.77%
Baby fed on regular schedule	0.861	16.03%	**t = 4**		
			Child been choosy with food	1.000	67.70%
t = 2			Child refuses the right food	0.944	60.27%
Child being choosy with food	1.000	51.73%	Child has eating routine	0.566	21.67%
Child refuses food	1.103	62.91%	**t = 5**		
Child has eating routine	0.786	31.92%	Child been choosy with food	1.000	69.51%
Difficulties feeding	0.955	47.14%	Child refuses the right food	1.040	75.23%
			Child has eating routine	0.643	28.70%

Table B7: Curative Health Investment: loadings and signal

Measurement	α	s_θ	Measurement	α	s_θ
t = 1			**t = 4**		
Ever called out doctor for baby	1.000	21.78%	Doctor has seen child at surgery	1.000	44.33%
No. of medications since at home	0.512	5.74%	Doctor called to home	0.637	17.75%
t = 2			Child had cough medicine	0.236	2.49%
Visits to doctor	1.000	34.74%	Child had antibiotics	0.725	22.93%
Doctor visits to home	0.936	29.66%	Child had diarrhea medicine	0.211	1.95%
Child had cough medicine	0.387	5.10%			
Child had antibiotics	0.812	22.75%	**t = 5**		
Child had diarrhea medicine	0.472	7.60%	Doctor has seen child at surgery	1.000	47.01%
t = 3			Doctor called to home	0.491	11.05%
Doctor has seen child at surgery	1.000	37.83%	Child had cough medicine	0.225	2.34%
Doctor called to home	0.789	24.79%	Child had antibiotics	0.763	26.75%
No. of doctor visits to home	0.752	22.61%	Child had diarrhea medicine	0.166	1.25%
Child had cough medicine	0.232	2.23%			
Child had antibiotics	0.677	18.26%			
Child had diarrhea medicine	0.208	1.75%			

Table B8: Mother Socio-emotional Factor: loadings and signal

Measurement	α	s_θ
Interpersonal awareness	1.000	77.15%
Need for approval	0.559	24.54%
Separation anxiety	0.896	62.10%
Timidity	0.738	42.19%
Fragile inner-self	0.850	55.84%

Table B9: Mother Health Factor: loadings and signal

Measurement	α	s_θ
Health up to present pregnancy	1.000	99.90%
Health History	0.190	3.62%
Recent health problem	0.189	3.54%
Problems requiring regular treatment	0.151	2.20%

Table B10: Initial Condition θ_0: loadings and signal

Measurement	α	s_θ
Birth weight	1.000	90.89%
Head circ.	0.691	49.67%
Crown-heel length	0.747	53.24%
Placental weight	0.678	40.12%

$M_t^{k,l}$ we are assuming that $\alpha_t^{k',l} = 0$ for all $k' \neq k$. Although quite stringent, this assumption facilitates the interpretation of the latent traits that we are extracting from the system: each factor represents only the variation that is common to its dedicated measurements, and those measurements alone.

However, identification requires much less stringent assumptions on the matrix $\mathbf{\Lambda}$, therefore we can test the validity of our restrictions. A first test can be performed by running multiple times the factor analysis, each time adding to the measurement system of factor θ_t^k one of the measures dedicated to the other factors $\theta_t^{k'}$, for $k' \neq k$. For example, when estimating the cognitive factor at time t, we add one-by-one all of the measures that are associated to the socio-emotional, health and investment factors at time t.

Collecting all of the estimated factor loadings $\alpha_t^{k'}$ and the associated commonalities $s_\theta^{t,k',l} = 1 - s_\varepsilon^{t,k',l}$, we see some patterns emerging[51]: first of all, the great majority of the factor loadings are very low, with an average of 0.19 and a median of 0.10, and the noise-to-signal ratio is very high, with an average uniqueness $s_\varepsilon^{t,k',l}$ of 97.8% and a median of 99.6%. It is easy to notice that the distribution of the estimated coefficients is quite skewed, and indeed most of the measurements associated to any other factor k' would not have been included in the measurement of factor k when following the procedure of measurement selection described in appendix B.B1.

However this is not the true for some particular cases: notably the measures related to parental investment θ_t^{PI} usually display a mild loading on the cognitive and, to a lesser extent, the socio-emotional factor, especially after period 2. The same is true for the measures of curative health care, which display high (negative) loadings on the health factor, and vice-versa. This should not come as a surprise, since mothers take their children to the doctor when they are sick and investment has a positive effect on skills, but both skills and investment are persistent over time: this feature of the data is already embedded in our model and is reflected in the estimated contemporaneous covariance between the factors as well as their relation over time via the technology. Furthermore, we want to make a clear distinction between parental choices, which are reflected in the investment factors, and characteristics of the children, which are captured by the skills factors. Therefore we believe that imposing a zero loading across the investment and the skills factors is an important feature of our analysis.

Focusing only on the cross-loadings within skills (cognitive, socio-emotional and health) and within investment (parental investment, preventive health care and curative health care), we find that the assumptions of dedicated measurement is reasonable. The average loadings $\alpha_t^{k'}$ for $k' \neq k$ is 0.12 and the median is 0.08, while the average uniqueness $s_\varepsilon^{t,k',l}$ is 99.1% and the median 99.7%.

A second investigation about the soundness of the hypothesis of dedicated measurement is to run a regression of each measure on all the factor scores - extracted using the current measurement system - in order to evaluate the correlation between each latent variable and the measurement:

$$\text{(B2)}\quad M_t^{k,l} = \kappa + \beta_t^{C,l}\theta_t^C + \beta_t^{SE,l}\theta_t^{SE} + \beta_t^{H,l}\theta_t^H + \beta_t^{PI,l}\theta_t^{PI} + \beta_t^{PH,l}\theta_t^{PH} + \beta_t^{CH,l}\theta_t^{CH} + \epsilon_t^{k,l}$$

When performing the regression shows in equation $(B2)$, we find a patter of results

[51]The 960×2 estimated parameters are not reported here for brevity. All are available from the author upon request.

very similar to the ones associated to the previous test. The average coefficient $\beta_t^{k',l}$ associated to a measure that has been dedicated to another factor ($k' \neq k$) is 0.060 and the median is 0.036. When focusing only on the cross loadings within skills and within investment, the average $\beta_t^{k',l}$ for $k' \neq k$ is 0.049 and the median in 0.037.

C. Estimation Procedure

The estimation procedure used throughout the paper is MLE based on the assumption of joint normality of error terms $\boldsymbol{\varepsilon}_t, \eta_t$ and latent factors $\boldsymbol{\theta}_t^k$. Other less efficient estimation procedures that we consider here below as robustness checks are regression on factor scores and regression on indexes.

C1. Regression on Factor Scores

A simpler estimation technique would be to use the so-called "regression on factor scores". This is a three-step procedure which can be considered more intuitive. In the first step, a factor model is *separately* estimated for each latent trait in each time period; then factor scores are *separately* predicted for each variable; finally, the third step consists in a ordinary least square analysis performed among the factor scores.

Although such multi-stage procedure is much more straightforward to estimate, the asymptotic and finite sample performance of such estimator are not known and could be potentially very poor. In a simple static model, Skrondal and Laake (2001) demonstrate that the conventional approach to factor score regression performs very badly; although they propose a "revised" approach, this revision is not suited for a dynamic model.

Nevertheless, we provide estimates using this procedure for comparison. Table (C1) show some differences in the point estimation of the structural parameters, and a substantial unreliability of the standard errors estimated using bootstrapping. The difference is more pronounced in the estimation of the effect of curative health care (parameters $delta_3^k$): this can be due to the fact that the multi-step procedure does not fully account for the dynamic structure of the model, or for the joint correlation between the latent traits.

D. Omitting the Health Factor

If we do not control for current health, but we still include the potential effect of health investment on the evolution of cognitive and non cognitive skills, we obtain the estimates presented in table (D1). We can see that indeed there is a sharp distinction between curative health care and preventive care: while the latter still has a positive effect on both cognitive and socio-emotional development of the child, the former seems to be detrimental to both once we fail to take into consideration the effect of current health.

Table C1: Estimating the Technology of Skill Formation Using Regression of Factor Scores

		(1) Cognitive Skills θ_{t+1}^{C}	(2) Socio-emotional Skills θ_{t+1}^{SE}	(3) Health θ_{t+1}^{H}
Cognitive Factor θ_t^C	γ_1	0.685	0.022	0.002
		(0.006)	(0.003)	(0.002)
Socio-emotional Factor θ_t^{SE}	γ_2	0.032	0.580	0.012
		(0.004)	(0.005)	(0.002)
Health Factor θ_t^H	γ_3	0.018	0.104	0.808
		(0.007)	(0.007)	(0.004)
Parenting Investment θ_t^{PI}	δ_1	0.096	0.085	0.005
		(0.006)	(0.005)	(0.002)
Preventive Care θ_t^{PH}	δ_2	0.003	0.050	0.016
		(0.005)	(0.003)	(0.002)
Curative Care θ_t^{CH}	δ_3	0.005	0.002	-0.134
		(0.007)	(0.006)	(0.004)
Mother Education θ_P^C	β_1	0.027	0.004	-0.011
		(0.002)	(0.002)	(0.001)
Mother Socio-emotional θ_P^{SE}	β_2	0.009	0.029	0.000
		(0.003)	(0.003)	(0.001)
Mother Health θ_P^H	β_3	-0.015	0.108	0.096
		(0.010)	(0.009)	(0.007)

Bootstrapped Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for the measurements used.

Table D1: Technology of Cognitive and Socio-emotional Skill Formation with the Investment in Health

		(1) Cognitive Skills θ_{t+1}^{C}	(2) Socio-emotional Skills θ_{t+1}^{SE}
Cognitive Factor θ_t^C	γ_1	0.634	0.010
		(0.002)	(0.005)
Socio-emotional Factor θ_t^{SE}	γ_2	0.038	0.478
		(0.004)	(0.005)
Parenting Investment θ_t^{PI}	δ_1	0.135	0.140
		(0.007)	(0.009)
Preventive Care θ_t^{PH}	δ_2	0.017	0.092
		(0.005)	(0.006)
Curative Care θ_t^{CH}	δ_3	-0.051	-0.112
		(0.007)	(0.008)
Mother Education θ_P^C	β_1	0.055	-0.012
		(0.002)	(0.003)
Mother Socio-emotional θ_P^{SE}	β_2	0.009	0.074
		(0.004)	(0.004)
Mother Health θ_P^H	β_3	0.003	0.036
		(0.003)	(0.004)

Standard errors in parentheses; ALSPAC children aged 0-5; Latent variables estimated using linear factor model. Controls X: constant, age of the child at the assessment date in months, gender dummy, parity. See tables $(1-4)$ for definition of factors

UBS International Center of Economics in Society
University of Zurich
Department of Economics
Schönberggasse 1
CH-8001 Zurich
Tel. +41 44 634 57 22
contact@ubscenter.uzh.ch
www.ubscenter.uzh.ch

Economics. *For Society.*