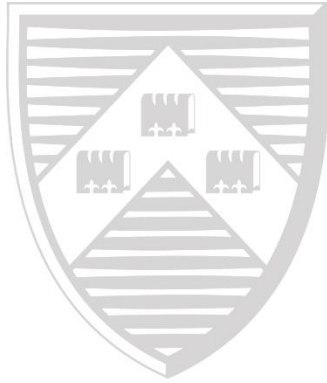


UNIVERSITY *of York*



Discussion Papers in Economics

No. 19/08

Child Socio-Emotional Skills:
The Role of Parental Inputs

Gloria Moroni, Cheti Nicoletti, Emma Tominey

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Child Socio-Emotional Skills: The Role of Parental Inputs*

Gloria Moroni[†] Cheti Nicoletti[‡] Emma Tominey[§]

June 10, 2019

Abstract

Informed by the psychological literature and our empirical evidence we provide new insights into the technology of socio-emotional skill formation in middle childhood. In line with economic evidence, increasing parental inputs that enrich the child home environment and reduce stress has larger returns for children with higher socio-emotional skills in early childhood (complementarity), but only for levels of inputs that are high. For low levels of inputs, i.e. levels implying a stressful home environment, an increase has a higher return for children with lower socio-emotional skills in early childhood (substitutability). Consequently, well targeted policies can reduce middle childhood socio-emotional gaps.

Keywords: Socio-emotional skills, Complementarities, Substitutabilities, Parenting styles, Mother’s mental health, Time investment, Child behavioural disorders, Diathesis-stress hypothesis.

JEL Classification: J13, D10, I10, I31

*We thank for comments and feedback the participants at the NCDS 60 years of our Lives in London (2018), NHH Norwegian School of Economics FAIR Seminar Series (2018), Royal Economic Society Junior Symposium in Brighton (2018), International Workshop on Applied Economics of Education in Catanzaro (2018), Summer School on Socioeconomic Inequality (HCEO) in Bonn (2018), Workshop on Labour and Family Economics in York (2018), European Association of Labour Economists in Lyon (2018), Royal Economic Society in Warwick (2019), Society for Labour Economists in Arlington (2019), Inequality, Social Origins and Intergenerational Mobility Workshop in Milan (2019). We would like to thank the Millennium Cohort Study (MCS) families, as well as the MCS team at the Institute of Education. This research was partly supported by the Economic and Social Research Council through their grants to the Research Centre on Micro-Social Change in ISER (ES/L009153/1).

[†]DERS, University of York. Corresponding author: gm889@york.ac.uk.

[‡]DERS, University of York; ISER, University of Essex, CHILD Collegio Carlo Alberto; IZA

[§]DERS, University of York; HCEO; IZA

1 Introduction

Evidence suggests that gaps in child socio-emotional skills emerge at very young ages, persist across the lifetime and are predictive of adult skills including labour market, health and crime outcomes.^{1,2} In this paper we assess whether it is possible to reduce the gap in socio-emotional skills in middle childhood by targeting interventions at specific parental inputs. We focus on the development from age 6 to 11 - a sensitive period, when socio-emotional skills are especially reactive to environmental stimuli (see Cunha and Heckman 2008), as well as a developmental switch point, when environmental inputs combine with child's predisposition to determine personality and physical attributes (see Del Giudice 2018).

The main body of the literature on the production of child cognitive skills has found empirical evidence for complementarity (a higher productivity of an input for children with high early cognitive skills) for almost all inputs.³ In the presence of a similar complementarity for the production of child socio-emotional skills in middle childhood, we would expect a higher return to parental inputs in middle childhood for children with higher socio-emotional skills in early childhood. Such complementarity is exemplified in Figure 1 by the solid regression lines for socio-emotional skills on parental inputs with a flatter slope for children with low (Panel A) rather than high (Panel B) socio-emotional skills in early childhood.

However, the development of child socio-emotional skills is expected to show substitutability - a higher productivity of an input for children with low early socio-emotional skills - if the child is exposed to a stressful home environment. This is because child socio-emotional and behavioural disorders are triggered by the interaction of a stressful environment and a predisposition to the disorder, as described by the *diathesis-stress* hypothesis in the psychological literature (see Monroe and Simons 1991; Belsky and Pluess 2009). Therefore, we expect children with low socio-emotional skills in early childhood, which can be an indicator for a predisposition to behavioural disorders, to be more vulnerable to a stressful home environment in middle childhood; and children with high socio-emotional skills in early childhood to be more resilient to these stressful experiences. Because low levels of parental inputs, such as sensitive parenting and time investments, are likely to cause a stressful home environment, while high levels are associated with an enriched home environment, we expect

¹What we call socio-emotional skills have been referred to as non-cognitive or soft skills (or abilities).

²See Heckman and Rubinstein (2001); Carneiro and Heckman (2003); Heckman et al. (2006); Kautz et al. (2014).

³Empirical evidence on the degree of complementarity between investments and cognitive skills in the production of future cognitive skills can be found in Cunha et al. (2010), Nicoletti and Rabe (2014), Attanasio et al. (2015, 2017), García and Gallegos (2017).

an increase in parental inputs when low, to have higher productivity for children with low socio-emotional skills. This pattern is described by the dotted lines in Figure 1. For low levels of parental inputs - or a stressful home environment - an improvement in the home environment has a greater effect for children with low rather than high socio-emotional skills in early childhood (compare Panels A and B). This is consistent with some degree of substitutability between low levels of parental inputs and early child socio-emotional skills and it implies that interventions in middle childhood to increase these inputs - i.e. to improve a stressful home environment - would be effective and optimally timed to facilitate the catch up of socio-emotional skills.

To check empirically whether there is evidence of substitutability rather than complementarity for low levels of parental inputs, we estimate a production model of socio-emotional skills at age 11 (middle childhood) that allows the productivity of a set of parental inputs to vary both across the distribution of child's socio-emotional skills at age 6 (early childhood) and across the input distribution. The only other papers which have allowed the productivity of inputs to vary across level of earlier socio-emotional skills are [Cunha et al. \(2010\)](#) and [Attanasio et al. \(2015\)](#). They find evidence for complementarity but under the assumption that the level of complementarity/substitutability does not change across different inputs and across different levels of the inputs.

The parental inputs measured between age 6 and 11 include (1) sensitive parenting, (2) routines in parenting, (3) parental time investment; (4) family income, (5) mother socio-emotional skills, (6) mother cognitive skills. We allow for self-productivity of socio-emotional skills and cross-productivity between the child cognitive and socio-emotional skills. Factor analysis is used to construct two measures - or traits - of socio-emotional skills, of internalising and externalising behaviours and we assume that both measures follow the same production model. Typical examples of internalising behaviours are being withdrawn, anxious, depressed and having somatic symptoms. Externalising behaviours occur when children direct their negative feelings and problems outward by adopting disruptive, hyperactive and aggressive behaviours.⁴ Children with high levels of either internalising or externalising problems, i.e. in the first quartile of the distribution, are 7 times more likely to develop Attention Deficit Hyperactivity Disorder (ADHD) and 20 times more likely to be diagnosed with Autism Spectrum Disorder (ASD).

⁴The concepts of internalising and externalising behaviours were introduced in the psychological literature by [Achenbach \(1966\)](#), who used factor analysis to show that these two socio-emotional traits are the two main latent factors underlying a large set of psychological disorders. Since then, the concepts of internalising and externalising behaviours have been used in more than 75,000 articles in psychology (see [Achenbach et al. 2016](#)).

Our first contribution is to adopt a flexible production model of socio-emotional skills which relaxes the assumption of an equal degree of complementarity/substitutability across inputs and across levels of inputs, therefore allowing to test whether the previous evidence of complementarity between early socio-emotional skills and parental inputs in middle childhood is rejected for some of the parental inputs. In particular, we test whether the complementarity assumption is rejected because children with low (high) levels of socio-emotional skills in early childhood are more vulnerable (resilient) to a stressful environment, i.e. are more vulnerable (resilient) to low levels of parental inputs that are stress relievers. Looking at Figure 1 Panels A and B, the complementarity hypothesis is represented by the solid lines, whereas the alternative hypothesis is represented by the dotted lines.

Our second contribution is to estimate the production of socio-emotional skills separately by gender, therefore providing new insights on gender differences in skills development during middle childhood.⁵ This contribution is especially important because middle childhood is characterised by the awakening of the adrenal glands which leads the brain to develop along different pathways for boys and girls (Del Giudice, 2018).

Our final contribution is to control for unobserved heterogeneity by adopting child fixed effect estimation which exploits the availability of two traits of socio-emotional skills for each child - internalising and externalising behaviours.⁶ Our identification strategy relies on variation in the deviation of a trait-specific socio-emotional skill from the mean *across* the two traits, i.e. the mean across externalising and internalising behaviours. Whilst variation in the general level of socio-emotional skills across individuals at age 6 is likely to be endogenous, variation in the deviation of a trait-specific skill from the mean is mainly driven by a predisposition to either internalising or externalising problems. The consistency of this child fixed effect estimation relies neither on the assumption that unobserved inputs be independent of observed inputs nor on the assumption that they be invariant across life stages, age or siblings; which are the most common assumptions imposed by previous papers on child development. We discuss below the three identification assumptions of our estimation method.

First, we assume that the latent externalising and internalising traits follow the same dynamic production model. Section 6 provides references to empirical papers with results supporting the assumption that the effect of parental inputs are similar for externalising and internalising behaviours. Furthermore, we show that our main results are robust to

⁵For previous evidence on gender difference, see Bertrand and Pan (2013); Baker and Milligan (2016).

⁶A similar approach is adopted by Dee 2007; Clotfelter et al. 2010; Bandiera et al. 2010; Lavy et al. 2012 using within-student between-subject estimation.

relaxing the assumption of invariance of the self-productivity coefficient across the two traits. Second, we assume that the variation picked up by the deviation of a trait-specific socio-emotional skill from the mean across the two traits is explained by random variation in the child predisposition to either internalising or externalising problems. Section 6 rules out alternative sources of variation between internalising and externalising which could be endogenous. Third, we assume there are no omitted inputs whose productivity varies across the distribution of early socio-emotional skills. Even if our estimation model controls for all unobserved inputs that are invariant across socio-emotional traits, we could still have a potential endogeneity of parental inputs if the productivity of these omitted inputs varies across different levels of socio-emotional skills in early childhood and are correlated with our parental inputs. For this reason, in all of our analysis we control for a large set of further characteristics which can proxy for unobserved inputs, including controls for mother’s labour supply, child care arrangements, fertility decisions, neighbourhood capital, social capital and health capital. Finally, we are also able to show that our results are unchanged even when controlling for past parental inputs.

Using the UK Millennium Cohort Study, we find that levels of sensitive parenting, parental time investments and maternal mental health below (above) the median constitute examples of a stressful (enriched) home environment. Children with low socio-emotional skills in early childhood are more vulnerable to a stressful home environment in middle childhood - defined as levels of these inputs below the median - whilst high skilled children in early childhood show resilience. Our empirical evidence supports the type of relationship between socio-emotional skills and parental inputs in middle childhood depicted in Figure 1 Panels (A) and (B) by the dotted lines for low levels of inputs and by the solid lines for high level of inputs. Therefore, our evidence suggests that there is complementarity between parental inputs and early socio-emotional skills, but only for high starting values of parental inputs, i.e. corresponding to an enriched environment. We instead find substitutability for low levels of parental inputs that constitute examples of a stressful environment. Across gender, boys with low (high) socio-emotional skills seem more vulnerable (resilient) to low levels of sensitive parenting; whereas girls seem to be vulnerable (resilient) also to low levels of parental time investments and maternal mental health. On the other hand we find no variation in the productivity across level of early socio-emotional skills in family income, routines parenting style, mother’s cognitive skills and child’s cognitive skills.

Taken all together, these findings suggest that there are some parental inputs which may help children to catch up their socio-emotional skills, and these seem to be the inputs that reduce child stress. A back of the envelope calculation shows that if interventions to

raise parent inputs were targeted correctly at households with low initial levels of the inputs displaying substitutability, then the gap between children at the bottom - and the top of socio-emotional skills distribution at age 11 could be reduced by up to 34% for boys and 52% for girls by increasing the input by one standard deviation.

The remainder of this paper is organized as follows. We introduce the production model and the estimation strategy in Sections 2 and ?? respectively. We then move to describe the data and variables in Section 4 and to present our main empirical results in Section 5. In Section 6 we test the validity of our model assumptions and in Section 7 present robustness checks to show that there is no bias caused by measurement error, model specification or the specific metric of socio-emotional skills. Finally, Section 8 concludes.

2 The production model of socio-emotional skills

In this section we lay out our production model that describes the process of development of children’s socio-emotional skills between age 6 and 11. We allow children’s socio-emotional skills at age 11 to depend on their past socio-emotional and cognitive skills and on their family inputs, as described by the following equation:

$$\theta_{i,s,2}^N = f_2\left(\theta_{i,s,1}^N, \theta_{i,1}^C, \mathbf{I}_{i,2}, \boldsymbol{\theta}_{P,i,2}, \mu_{i,2}^N\right) \quad (1)$$

where the subscripts 1 and 2 denote the development stages from 0 to 6 and from 6 to 11; s takes value 1 for internalising and 2 for externalising behaviour; $\theta_{i,s,1}^N$ and $\theta_{i,s,2}^N$ are the unobserved latent trait-specific socio-emotional (non-cognitive) skills of the child i for the trait s at the end of stage 1 and 2, i.e. at age 6 and 11; $\theta_{i,1}^C$ is the child’s unobserved latent cognitive skill at the end of stage 1, i.e. at age 6. $\mathbf{I}_{i,2}$ and $\boldsymbol{\theta}_{P,i,2}$ represent two vectors of parental inputs, which are parental investments and skills during stage 2, i.e. between age 6 and 11.⁷ $\mathbf{I}_{i,2} = [I_{i,2}^{Sensitive}, I_{i,2}^{Routines}, I_{i,2}^{Time}, I_{i,2}^{Income}]$ is a row vector of four parental investments which are measures of sensitive parenting style, daily routine parenting style, parental time investment and family income; while $\boldsymbol{\theta}_{P,i,2} = [\theta_{P,i,2}^{Cognitive}, \theta_{P,i,2}^{Socio-emotional}]$ is a row vector of parental skills, where mother’s cognitive skills are proxied by mother’s education and mother’s socio-emotional skills are proxied by mother’s mental health. $\mu_{i,2}^N$ denotes potential omitted inputs that might affect the socio-emotional development of the child during stage

⁷We cannot observe these parental inputs for the full period from 6 and 11 and we assume that measures observed when the child is 7 are a good proxy for the parental inputs during stage 2.

2. We do not restrict the unobserved component $\mu_{i,2}^N$ and the production function f_2 to be time invariant across different child's life stages, as done in some of previous papers; but we assume invariance of f_2 and $\mu_{i,2}^N$ between the two trait-specific socio-emotional skills.⁸

While the end of stage 1 is set at age 6, the start does not need to be at age 0 and we could divide the development stage from 0 to 6 in different shorter stages, e.g. from 0 to 1, 1 to 2, 2 to 4 and 4 to 6, without any loss of generality for our production model. For all these earlier development stages, we assume that child socio-emotional skills evolve following a stage-specific dynamic production function, say $\theta_{i,s,t}^N = f_t()$ for each stage $t < 2$, and that all these production functions are twice continuously differentiable in their inputs.

Our aim is to assess whether there are interaction effects between each of the parental inputs in $[\mathbf{I}_{i,2}, \boldsymbol{\theta}_{P,i,2}]$ and the socio-emotional skill $\theta_{i,s,1}^N$, i.e. to test whether the productivity of each of the seven parental inputs varies across levels of the socio-emotional skill at age 6. As highlighted above, there is almost no empirical evidence on the presence of such interactions in the production of socio-emotional skills. The presence of complementarity (a positive interaction effect) between a parental input in stage 2 and the socio-emotional skill at the end of stage 1 would imply the presence of dynamic complementarity between the parental inputs in stage 2 and inputs in earlier stages (see Heckman and Mosso 2014).⁹ On the contrary, the presence of substitutability (a negative interaction effect) between a parental input in stage 2 and the socio-emotional skill at the end of stage 1 would imply dynamic substitutability between such input in stage 2 and earlier stages' inputs. Knowing whether there exists substitutability or complementarity is therefore relevant to understand which inputs in stage 2 are most effective to reduce potential gaps in socio-emotional skills.

We allow the production function (1) to be flexible enough to accommodate for the

⁸We provide empirical evidence to support the model invariance across traits in the empirical section.

⁹We can show this by measuring the degree of dynamic complementarity between $I_{i,t}^{Time}$ and $I_{i,t-\delta}^{Time}$ for the production of socio-emotional skills at the end of stage t , $\theta_{i,s,t}^N$, in the following way,

$$\frac{d^2 f_t()}{dI_{i,t}^{Time} dI_{i,t-\delta}^{Time}} = \frac{\partial^2 f_t()}{\partial I_{i,t}^{Time} \partial \theta_{i,s,t-1}^N} \left(\prod_{j=1}^{\delta} \frac{\partial f_{t-\delta+j}()}{\partial \theta_{i,s,t-\delta+j-1}^N} \right) \frac{\partial f_{t-\delta}()}{\partial I_{i,t-\delta}^{Time}}. \quad (2)$$

Because $\frac{\partial f_{t-\delta+j}()}{\partial \theta_{i,s,t-\delta+j-1}^N}$ and $\frac{\partial f_{t-\delta}()}{\partial I_{i,t-\delta}^{Time}}$ are generally higher than zero, a strong static complementarity between $I_{i,t}^{Time}$ and $\theta_{1,s,i,t-1}^N$, i.e. a large value for $\frac{\partial^2 f_t()}{\partial I_{i,t}^{Time} \partial \theta_{i,s,t-1}^N}$, will lead to a strong dynamic complementarity between $I_{i,t}^{Time}$ and $I_{i,t-\delta}^{Time}$. Following a similar reasoning we can show also that any early input in $t - \delta$ is complement with a later input in t as long as there is complementarity between this later input in t and the socio-emotional skill in $(t - 1)$.

presence of interactions between each pair of inputs in the following way:

$$\begin{aligned} \theta_{i,s,2}^N = & m_{2,N}(\theta_{i,s,1}^N) + \sum_{k=1}^K m_{2,N,k}(\theta_{i,s,1}^N, Z_{k,i,2}) + \sum_{k=1}^K m_{2,k}(Z_{k,i,2}) + \sum_{k=1}^K \sum_{\substack{j=2 \\ j>k}}^K m_{2,k,j}(Z_{k,i,2}, Z_{j,i,2}) \\ & + m_{2,\mu}(\mu_{i,2}^N) + m_{2,N,\mu}(\theta_{i,s,1}^N, \mu_{i,2}^N) + \sum_{k=1}^K m_{2,k,\mu}(Z_{k,i,2}, \mu_{i,2}^N), \end{aligned} \quad (3)$$

where $k = 1, \dots, K$, $Z_{k,it}$ denotes the k -th input in the following vector of dimension $K \times 1$ and that we express in deviation from its sample mean,

$$\mathbf{Z}_{k,i,2} = \left(\theta_{i,2}^C, \mathbf{I}_{i,2}, \boldsymbol{\theta}_{P,i,2} \right).$$

The term $m_{2,N}(\theta_{i,s,1}^N)$ characterizes the effect of the lagged latent socio-emotional skill when all other inputs are at the mean, i.e. when $Z_{k,i,2} = 0$ for each k ; and $m_{2,k}(Z_{k,i,2})$ denotes the effect of the input $Z_{k,i,2}$ when $\theta_{i,s,1}^N$ is at the baseline value¹⁰ and all other variables are at the mean. $m_{2,N,k}(\theta_{i,s,1}^N, Z_{k,i,2})$ captures the differential productivity of each input $Z_{k,i,2}$ across levels of the lagged latent socio-emotional skill; $m_{2,k,j}(Z_{k,i,2}, Z_{j,i,2})$ measures the differential productivity of $Z_{k,i,2}$ at different values of the j -th input $Z_{j,it}$; and finally $m_{2,\mu}(\mu_{i,2}^N)$, $m_{2,N,\mu}(\theta_{i,s,1}^N, \mu_{i,2}^N)$ and $m_{2,k,\mu}(Z_{k,i,2}, \mu_{i,2}^N)$ are the effects of unobserved inputs and of their interactions with lagged socio-emotional skills and inputs.

Similar types of child development production functions with interactions between pairs of inputs have been considered by [Liu et al. \(2010\)](#), [Mueller \(2013\)](#) and [Nicoletti and Rabe \(2017\)](#) and are more general than the constant elasticity of substitution (CES) and Cobb-Douglas functions (see e.g. [Cunha et al. 2010](#); [Del Boca et al. 2013](#)). The advantage of adopting such general production function is that we can allow for self-productivity of socio-emotional skills, for cross-productivity between cognitive and socio-emotional skills and for complementarity/substitutability between each pair of inputs and between each of the inputs $Z_{k,i,2}$ and the socio-emotional skill $\theta_{i,s,1}^N$ without imposing the elasticity of substitution to be the same for all pairs of inputs as imposed by CES models.

¹⁰In our empirical application we consider dummies for the different quartiles of the lagged socio-emotional skill, $\theta_{i,s,1}^N$, where the baseline category is the first quartile.

3 Estimation of the production model

To estimate consistently the production model we need to take account of potential endogeneity caused by the omission of variables and of measurement error issues. The rest of this section is organized as follow. Section 3.1 introduces our estimation procedure of the production model and explains how we deal with endogeneity issues under the assumption of no measurement errors. In Section 3.2 we describe the measurement model of socio-emotional skills and parental inputs, which we use to derive the latent factors that are theoretically unaffected by measurement errors.

3.1 Estimation

Thanks to the assumption of invariance of the production model (3) across the two trait-specific socio-emotional skills, internalising and externalising behaviours, we can exploit the availability of the two trait-specific socio-emotional skills for each child to apply a child fixed-effect estimation.¹¹ We do this by rewriting the equation (3) in deviations from the mean across traits,

$$\ddot{\theta}_{i,s,2}^N = \ddot{m}_{2,N}(\theta_{i,s,1}^N) + \sum_{k=1}^K \ddot{m}_{2,N,k}(\theta_{i,s,1}^N, Z_{k,i,2}) + \ddot{m}_{2,N\mu}(\theta_{i,s,1}^N, \mu_{i,2}^N) + \ddot{u}_{i,s,2}, \quad (4)$$

where the double dot denotes the deviation from the mean across the two different traits, internalising and externalising. This transformation of $\theta_{i,s,2}^N$ cancels out any potential measurement error in the socio-emotional skills which is shared across traits and likewise any shared error in the lagged socio-emotional skills $\theta_{i,s,1}^N$.¹² This deviation transformation eliminates all the terms in the right hand side of the production model (3) that do not vary across the two trait-specific socio-emotional skills. Notice that the last two terms in equation 3 depends on the unobservable component $\mu_{i,2}^N$ and eliminating them solves, at least in part, the endogeneity caused by unobservables.

Computing the deviations from the mean across traits does not solve completely the endogeneity issue caused by unobservables $\mu_{i,2}^N$ because it does not eliminate the interactions between the child socio-emotional skills at age 6 and the unobservable parental inputs:

¹¹Notice that this child fixed effect allows to take into account of potential cross-productivity between internalising and externalising traits at each stage.

¹²If the production model is not linear in $\theta_{i,s,1}^N$, then considering a Taylor approximation of order one of the non-linear function of $\theta_{i,s,1}^N$ around the average of $\theta_{i,s,1}^N$ across traits would still cancel out the shared measurement error through the deviation from the mean transformation.

$m_{2,S,\mu}(\theta_{i,s,1}^N, \mu_{i,2}^N)$. These unobserved inputs bias our estimation of the productivity of observed parental inputs, $[\mathbf{I}_{i,2}, \boldsymbol{\theta}_{P,i,2}]$, only if both of the following conditions hold: (i) the productivity of the omitted inputs varies across the distribution of socio-emotional skills at age 6; (ii) there is correlation between the observed parental inputs and the omitted inputs. Previous papers on socio-emotional development have not considered the possibility of such bias. On the contrary, we control for this potential bias by including the interactions between $\theta_{i,s,1}^N$ and a rich set of additional observed inputs, which can proxy for the unobserved inputs. In particular, we consider measured of neighbourhood capital, social capital, health capital, mother's labour supply, child care arrangements and fertility decisions. We denote each of these additional inputs with $W_{g,i,2}$, where $g = 1, \dots, G$ denotes each of the G inputs, and we assume that the interaction term between the lagged socio-emotional skills and the unobservable component $\mu_{i,2}^N$, $m_{2,S,\mu}(\theta_{i,s,1}^N, \mu_{i,2}^N)$, can be replaced by a set of interactions between the lagged socio-emotional skills and each of the additional variables $W_{g,i,2}$, $m_{2,N,g}(\theta_{i,s,1}^N, W_{g,i,2})$, so that our production model expressed in deviations from the mean across traits becomes

$$\ddot{\theta}_{i,s,2}^N = \ddot{m}_{2,N}(\theta_{i,s,1}^N) + \sum_{k=1}^K \ddot{m}_{2,N,k}(\theta_{i,s,1}^N, Z_{k,i,2}) + \sum_{g=1}^G \ddot{m}_{2,N,g}(\theta_{i,s,1}^N, W_{g,i,2}) + \ddot{u}_{i,s,2}. \quad (5)$$

In the following we impose some parametric assumptions on model (5), but we keep the specification flexible enough to allow for a differential productivity of the inputs across different levels of $\theta_{i,s,1}^N$. Specifically, we model socio-emotional skills in stage 2 as a linear function of socio-emotional skills in stage 1 and of all possible pairwise interactions between each of the inputs and quartiles of socio-emotional skills in stage 1 - with each of these interactions allowed to change for levels of parental inputs above and below the median. The model specification that we adopt is the following:

$$\ddot{\theta}_{i,s,2}^N = \alpha + \ddot{\theta}_{i,s,1}^N \rho + \sum_{k=1}^K \sum_{q=2}^4 \sum_{p=0}^1 [\ddot{D}_{i,s,q} D_{i,k,p} Z_{k,i,2} \beta_{k,q,p}] + \sum_{g=1}^G \sum_{q=2}^4 [\ddot{D}_{i,s,q} W_{g,i,2} \gamma_{g,q}] + \ddot{u}_{i,s,2}, \quad (6)$$

where the double dot still denotes the deviation from the mean across the two different socio-emotional traits; $D_{i,s,q}$ is a dummy taking value 1 if the observed $\theta_{i,s,1}^N$ lies between the $(q-1)$ -th and q -th quartile and 0 otherwise, and $\ddot{D}_{i,s,q}$ is its deviation from the mean across traits; $D_{i,k,0}$ is a dummy variable taking value 1 for values of $Z_{k,i,2}$ below its median, while $D_{i,k,1} = 1 - D_{i,k,0}$ is a dummy variable taking value 1 for values of $Z_{k,i,2}$ above its median; $Z_{k,i,2}$ and $W_{g,i,2}$ are measured in deviations from their sample mean; α is the intercept; ρ is a scalar parameter measuring the self-productivity (persistence) of socio-emotional skills, i.e. the effect of an increase in $\theta_{i,s,1}^N$ while all other inputs are held at their mean; $\beta_{k,q,p}$ is a

scalar parameter measuring the differential productivity of input $Z_{k,i,2}$ when children have lagged socio-emotional skills in the q -th quarter rather than in the first (baseline) quarter of the distribution and have values of $Z_{k,i,2}$ below its median if $p = 0$ and above it if $p = 1$, while all other inputs are held constant at their mean; finally $\gamma_{g,q}$ is a scalar parameter measuring the differential productivity of $W_{g,i,2}$ for children in the q -th quarter of the distribution of $\theta_{i,s,1}^N$ with respect to the bottom quarter, while all other inputs are held at their mean. In Section 6 and 7 we allow for even more flexible specifications of the production model and provide empirical evidence that the parametric assumptions imposed in equation (6) are not too restrictive.

Estimating equation (6), which is expressed in deviations from the mean across traits, is equivalent to adopting a child fixed effect estimation. Similar types of child fixed effect estimations have been used in several previous papers, e.g. to estimate the effect of lagged cognitive abilities or school class characteristics and school peers on cognitive abilities or educational achievements.¹³ The main difference is that whilst we exploit variation within child across socio-emotional traits to estimate a production model for socio-emotional skills, previous papers exploit variation within child across cognitive skills such as subject-specific test scores. Furthermore, we allow for complementarities between inputs and lagged socio-emotional skills, similarly to Nicoletti and Rabe (2014) who consider complementarities between school investments and cognitive ability at the end of primary school in the production of cognitive abilities in secondary school.

Differently from previous papers that have used dynamic factor models to estimate the production function of cognitive and socio-emotional skills (e.g. Cunha et al. 2010; Attanasio et al. 2015; Agostinelli and Wiswall 2016),¹⁴ we identify the effect of lagged socio-emotional skills in the production model by using exclusively variation coming from the differences between trait-specific latent socio-emotional skills, i.e. differences between $\theta_{i,s,1}^N$ and $\theta_{i,s',1}^N$ for $s \neq s'$, rather than variation coming from the general latent socio-emotional skill. This matters because, while variation in the general level of socio-emotional skills across individuals at age 6 is likely to be endogenous, variation in $\theta_{i,s,1}^N$ across traits is mainly explained by variation in an innate predisposition to specific socio-emotional issues.¹⁵

Part of the variation in $\theta_{i,s,1}^N$ across traits could be related to trait-specific unobserved

¹³E.g. Dee 2007; Clotfelter et al. 2010; Bandiera et al. 2010; Lavy et al. 2012; Del Boca et al. 2017; Nicoletti and Rabe 2017

¹⁴Note that the number of measures required to estimate equation (6) in dynamic factor models would increase exponentially for each interaction included, making it nearly impossible to estimate.

¹⁵We explore this assumption further in Section 6.

inputs and therefore potentially endogenous. In particular, children may learn or imitate internalising and externalising socio-emotional behaviours from parents based on their trait-specific socio-emotional skills. To check if such an endogeneity issue could be a concern in our estimation, we run some robustness checks by including as further inputs measures of trait-specific socio-emotional skills for the parents, with results that confirm our main conclusions.

3.2 Measurement models of socio-emotional skills and parental inputs

We have multiple measures for each of the two socio-emotional traits - internalising and externalising behaviours. We allow each of the observed measures of socio-emotional skills to be affected by two types of measurement error, a measurement error which is shared across the two *traits* and across all *measures*, and a measurement error which is instead trait- and measure- specific.

Because the observed traits of internalising and externalising behaviours are derived from the Strength and Difficulty Questionnaire which is administered to mothers, we could have a systematic error shared across all observed socio-emotional traits if mothers systematically under-report or over-report behavioural issues of their child. To allow for such shared measurement error across different observed traits of socio-emotional skills, we allow the observed trait-specific socio-emotional skill $\theta_{i,s,t}^N$ to depend on a shared measurement error in the following way:

$$\theta_{i,s,t}^N = \vartheta_{i,s,t}^N + v_{i,t}, \quad (7)$$

where $\vartheta_{i,s,t}^N$ is the true trait-specific latent skill and $v_{i,t}$ is the measurement error which is common across socio-emotional traits in stage t ($t = 1, 2$). Note that while the variation in $\theta_{i,s,t}^N$ is in part spurious and caused by the measurement error $v_{i,t}$; the variation of $\vartheta_{i,s,t}^N$ within child and across the two traits does not depend on $v_{i,t}$. In our estimation model we correct for the shared error $v_{i,t}$ by considering the deviations of $\theta_{i,s,t}^N$ from its mean across traits.

A second type of measurement error is within traits and across *measures*. We can deal with this because for each trait (internalising and externalising behaviour) we can observe multiple measures at the end of stage t , i.e. at age 6 for $t = 1$ and at age 11 for $t = 2$. Following previous papers on the technology of skill formation (see [Cunha et al. 2010](#) and [Agostinelli and Wiswall 2016](#)), we assume the following relationship between each of the observed measures of a trait-specific socio-emotional skills and the unobserved latent trait-

specific socio-emotional skill $\theta_{i,s,t}^N$:

$$Y_{i,s,t}^{N,m} = \lambda_{s,t}^m + \alpha_{s,t}^m \theta_{i,s,t}^N + e_{i,s,t}^m, \quad (8)$$

where $Y_{i,s,t}^{N,m}$ is the m -th observed measure of the s trait-specific socio-emotional skill at the end of stage t (at 6 or at 11), $m = 1, \dots, M_{s,t}$, $s = 1$ for internalising and 2 for externalising, $\lambda_{s,t}^m$ and $\alpha_{s,t}^m$ are the intercepts and factor loadings specific for the measure m . $e_{i,s,t}^m$ is the measure-specific measurement error which has mean zero and is assumed to be uncorrelated with $\theta_{i,s,t}^N$ and independently distributed across children and measures. The latent trait-specific socio-emotional skill $\theta_{i,s,t}^N$ can be identified only up to scale and location parameters. By setting the mean of $\theta_{i,s,t}^N$ to 0 and $\alpha_{s,t}^1 = 1$, we fix the location and scale so that we can identify the parameters in $\lambda_{s,t}^m$ and $\alpha_{s,t}^m$ for any trait s and t and we can extract the latent factor. We use measures of peer problems and emotional symptoms for internalising behaviours, and measures of hyperactivity and conduct problems for externalising behaviours. We use these measures observed at 5 and 7 as proxies for measures at 6; whereas we use only measures observed at age 11 as proxies for internalising and externalising at 11.¹⁶

We consider measurement models similar to (8) for the following parental inputs: sensitive parenting style, routine parenting style, parental time investment and mother's mental health.¹⁷ For each of these parental inputs which we denote with the subscript h , we consider the measurement model

$$I_{i,h,2}^m = \lambda_{h,2}^{I,m} + \alpha_{h,2}^{I,m} I_{i,h,2} + e_{i,h,2}^{I,m}, \quad (9)$$

where $I_{i,h,2}^m$ is the m -th observed measure of the h parental input at end of stage 2; $m = 1, \dots, M_{h,2}$; $h = 1$ for sensitive parenting style, 2 for routine parenting style, 3 for parental time investment and 4 for mother's mental health; $\lambda_{h,2}^{I,m}$ and $\alpha_{h,2}^{I,m}$ are the intercepts and factor loadings specific for the measure m ; $e_{i,h,2}^{I,m}$ is the measure-specific measurement error which has mean zero and is assumed to be uncorrelated with $I_{i,h,2}$ and independently distributed across children, measures and inputs.¹⁸ We identify and extract each of the latent factor $I_{i,h,2}$ by setting the factor mean to 0 and $\alpha_{h,2}^{I,1}$ to 1.

We estimate the production model by using directly the factors $\theta_{i,s,1}^N$, $\theta_{i,s,2}^N$ and $I_{i,h,2}$ for $h = 1, \dots, 4$, which are theoretically free of measurement error when expressed as deviations from the mean across traits.

¹⁶See Section 4 for more details.

¹⁷Notice that we adopt the same measurement model for child cognitive skills.

¹⁸Two parental inputs, family income and mothers' education, are directly used in our estimation without considering any factor model.

4 Description of data

We use the UK Millennium Cohort Study (MCS), a cohort member longitudinal survey covering a sample of about 19000 children born in the UK between September 2000 and January 2002. Interviews on the children and the family members are collected when the children are aged 9 months, 3, 5, 7, 11 and 14 years old.¹⁹ The MCS includes a wealth of information about the child, his/her household members and the home environment.

The sample consists of singleton children whose parents are married or cohabiting with analysis restricted to children for whom we observe socio-emotional skills reported by the mother, cognitive skills, parental inputs and parental skills. Our final sample consists of 6452 children. Table A.1 shows the demographic characteristics of the sample of children at age 7. The average age of the child in the sample is 86.6 months, with an average number of siblings of 1.5 and the mother being 37 years old. 93% of children are white; 36% of children have mothers in a managerial or professional occupation, 29% in an intermediate occupation and 32% in a routine or manual occupation.

4.1 Child's skills

4.1.1 Socio-emotional skills

We use the Strengths and Difficulties Questionnaire (SDQ) to construct our measures of socio-emotional skills. The SDQ is an international standardised test measuring children's behaviours and emotions in several contexts (Goodman, 1997). It is largely used in psychology as well as in the human capital literature within economics,²⁰ and it shows correlations with Diagnosis of Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD). The questionnaire, administered to parents since age 3 of the child, consists of 25 items on child psychological attributes measuring five scales for (i) Emotional Problems, (ii) Conduct Problems, (iii) Hyperactivity, (iv) Peer Relationship Problems and (v) Pro-social Behaviour are created. The scales can be either used separately to inform about specific behavioural problems or combined to create two different traits of socio-emotional skills such as internalising and externalising behaviour.

The internalising trait combines emotional symptoms and peer problems and the exter-

¹⁹More information on survey design, recruitment process and fieldwork can be found in Dex and Joshi (2005).

²⁰See for examples in the UK setting Del Bono et al. (2016); Moroni (2018).

nalising trait combines conduct problems and hyperactivity problems. Table 1 details the items included in each of the traits used in our analysis.²¹ The top panel of Table 2 reports the descriptive statistics of the four scales of socio-emotional skills of the child at age 5, 7 and 11 respectively for boys and girls. With the exception of emotional symptoms measured at age 5 and 7, girls have statistically significantly higher socio-emotional skills and are less likely to have emotional, peer, conduct or hyperactivity problems. This gender gap is more pronounced for externalising behaviour and especially for hyperactivity problems.

As explained in Section 3.2, we use factor models to derive the two specific traits of internalising and externalising behaviour. The correspondent factor loadings of the latent factors, i.e. internalising and externalising traits, can be found in Tables A.2 and A.3 in the Appendix, respectively for boys and girls.

4.1.2 Cognitive skills

Cognitive abilities are mainly derived from the British Ability Scales (BAS), which tests child skills differently depending on the age of the child (Elliott et al., 1997; Hansen et al., 2012) and captures child abilities including problem solving, expressive verbal ability and reading and spatial awareness.²²

The bottom panel of Table 2 reports the descriptive statistics of the different measures of child’s cognitive skills at age 5 and 7 respectively for boys and girls. Girls have statistically significantly higher cognitive skills than boys for all the measures, with the exception of reading abilities measured at age 5 where there is no statistically significant gender difference and maths at age 7 where boys outperform girls. Similarly to socio-emotional skills and as explained in Section 3.2, we use factor models to derive the latent cognitive skills of children. The correspondent factor loadings of the latent factors, can be found in Tables A.4 and A.5 in the Appendix for boys and girls respectively.

4.2 Main Inputs

The importance of considering a multiplicity of parental inputs when estimating the substitutability and/or complementarity between the inputs and the skills has been recently recognized by García and Gallegos (2017). We consider six main parental inputs grouped

²¹Pro-social behaviour is excluded from our analysis as it represents a different dimension of socio-emotional skills which cannot be combined with internalising or externalising behaviour.

²²More details on the tests administered to the child can be found in Hansen et al. (2012).

in two categories: parental investments and parental skills. We capture parental investments using measures of sensitive parenting style, routines parenting style, time investments and family income. We capture parental skills using measures of mothers’ cognitive and socio-emotional skills.

4.2.1 Parental Investments

The different types of parental investment included in our analysis capture various facets of parent-child interactions. Parenting style has recently been recognised as a distinct parental input from the usual material and time investments (Cobb-Clark et al., 2016). Mainly in the psychology literature, but recently also in economics (Doepke and Zilibotti, 2017; Fiorini and Keane, 2014; Cobb-Clark et al., 2016), different categories of parenting style have been used to measure how warm/communicative and how strict/controlling a parent is. We construct two domains of parenting styles capturing respectively the degree of sensitive parenting style and routines in parenting style.

The degree of sensitive parenting style is measured using the Straus’ Conflict Tactics Scale included in the MCS (Straus and Hamby, 1997). The mother reports on a 5 point scale ranging from ‘never’ to ‘daily’ how frequently she takes a set of specific actions with her child which are described in Table A.6.

We use latent factor models as explained in Section 3.2 to estimate the latent factor measuring sensitive parenting style which captures how harsh-sensitive a parent is when the child misbehaves. The latent variable for sensitive parenting style is reverse coded such that it can be interpreted as a positive input going from harsh to sensitive parenting, with higher values associated with more sensitive parenting. Table A.6 shows the factor loadings (Column 1) and the signal, i.e. percentage of the variance explained by the latent factor (Column 2). There is considerable variation in the amount of information captured by each measure of the same factor ranging from 11% to 50%.

The second domain of parenting style, i.e. routines in parenting style, accounts for the parental rules based on daily routines. The survey includes four questions on daily routines set by the parent on bedtime and watching TV, detailed in Table A.6.²³ The factor loadings in Table A.6 for the four measures (Column 1) and the signal (Column 2) are reported for the latent factor of routine parenting. We interpret this factor such that, the higher the value the stricter the daily routines.

²³Similar measures of parental rules have been used by Del Bono et al. (2016); Fiorini and Keane (2014).

The third type of parental investment considered in our analysis aims at capturing the parental time investment by using 6 questions regarding how often the mother spends time in formative activities with the child, again detailed in Table A.6. For all the statements the mother reports how frequently she plays these activities with the child in a scale ranging from ‘Every day or almost every day’ to ‘Not at all’.²⁴ The estimated factor loadings and signal are reported in Table A.6. The interpretation of the factor is such that the higher the value, the higher the time investment. For all factors in Table A.6, no measure has a signal close to 100% emphasizing the importance of accounting for measurement error through the latent factor model.

Finally, as largely established in the economic literature, income matters for the development of child human capital (see e.g. Dahl and Lochner 2012; Carneiro et al. 2015). We therefore include a fourth type of parental input to capture the material investment, measured with the household weekly income equivalised using the OECD scale²⁵ reported in the MCS.

4.2.2 Parental Skills

The second group of parental inputs consists of parental skills. We include mother cognitive skills, proxied by mother’s education (the age mothers left school) and mother’s socio-emotional skills captured by mother’s mental health. Specifically, mother’s mental health is assessed with the Mental Health – Kessler (K6) Scale (Kessler et al., 2002) a questionnaire intended to collect information on mothers’ psychological distress. The mother indicates how frequently she experiences feelings such as being depressed and hopeless for example, in the last 30 days ranging from ‘All of the time’ and ‘None of the time’. Similarly to the procedure followed for the other parental investments, we use latent factor model as described in Section 3.2 to estimate the latent factor measuring mother’s mental health. Table A.7 reports the factor loadings for the full set of measures (Column 1) and the signal, i.e. percentage of the variance explained by the latent factor (Column 2).

As additional measures of parental skills we include trait specific parents’ socio-emotional skills, of both mothers’ and fathers’. These measures will be used in our sensitivity analysis to show that the intergenerational transmission of trait-specific socio-emotional skills does

²⁴Similar measures of parental time investment have been used by Del Bono et al. (2016).

²⁵The OECD equivalised income is defined as the household income divided by the OECD equivalent scale, which counts the first adult as 1, each additional adult as 0.5 and each child (aged under 15) as 0.3 equivalent adults.

not represent a threat for our identification strategy. They are derived using the questions on parental neuroticism and extroversion available in the MCS and used to construct measures which proxy for parental internalising and externalising traits. Table A.8 reports the corresponding factor loadings (Column 1) and the signal (Column 2) for both the mother and the father.

4.3 Other inputs

Our production model of socio-emotional skills accounts for the fact that the child is exposed to other inputs, coming not only from the family environment but also from the external environment. We include in our production model of socio-emotional skills, a number of additional inputs such as health capital (including mother’s general health, reported by the mother as excellent, good, fair or poor, and indicators of child and father’s long-standing illness), neighbourhood capital (proxied by the local income decile) and social capital (measured as the presence of family and/or friends nearby). Finally, we include as additional controls other family decisions that might have been taken jointly with the parental investment decisions. These include (i) mother’s labour supply, measured as the number of hours worked; (ii) child care, measured by the number of daily hours the child is looked after not by parents (e.g. nursery, child minder, nanny, grandparents) and the number of daily hours the child spent in breakfast or afternoon club, (iii) mother’s fertility, i.e. whether the mother has given birth between age 5 and 7 of the cohort child.

4.4 Descriptive statistics of the inputs

Table 3 details the mean and standard deviation by gender of the inputs described above, all measured when the child is 7 years old. All the latent inputs that have been estimated with factor models - sensitive parenting style, routines in parenting style, time investment, mothers’ mental health and parents’ socio-emotional skills - have been standardized to have mean 0 and standard deviation of 1. The descriptive statistics show that there exist a statistically significant difference between boys and girls in the parental inputs that they receive. Parents adopt a harsher (less sensitive) parenting style with boys than with girls and invest more time in girls than in boys. On the contrary, there is no statistical gender difference in parental daily routines, mother’s mental health and parental socio-emotional skills. On average family OECD equivalised income is £497 and mothers are on average 18 years old when they leave education, with no differences between boys and girls.

Looking at the descriptives of the other inputs, 90% of the mothers report to have either excellent or good health, whilst 20%, 15% and 23-24% of boys, girls and fathers report a long-standing illness respectively. The areas where children live are on average ranked between the fifth and the sixth income decile (neighbourhood capital) and 69% of children live near family and friends (social capital). Finally, on average boys and girls spend 4.6 and 4.8 hours per day looked after by someone that is not the parent and they spend half an hour per day in either breakfast or afternoon club. The average number of hours worked by the mother is 17.4 per week and 13% or 14% of mothers has had another child when the cohort child was between 5 and 7 years old.

5 Estimation Results

We estimate the production model for socio-emotional skills at age 11 as described in equation (6) using child fixed effect estimation. The dependent variable $\theta_{i,s,2}^N$ is the trait-specific socio-emotional skill of the child at 11, in middle childhood (stage 2), which is a measure of internalising behaviour for $s = 1$ and of externalising behaviour for $s = 2$.²⁶ The right hand side variables are the trait-specific socio-emotional skills of the child at age 6, in early childhood (stage 1), $\theta_{i,s,1}^N$; the child cognitive skills in early childhood (age 6); the main parental inputs of interest during middle childhood, between age 6 and 11, which are sensitive parenting style, routines in parenting style, time investment, family income, mother's education and mothers' mental health, and a set of other controls which we use as proxies for potential omitted inputs to control for potential endogeneity. These extra control variables are mother working hours per week, hours spent by the child in non-parental care and in school clubs, fertility (the birth of another sibling), and three measures of neighbourhood, social and health capital (see Section 4.3 for more details).

Before presenting our results, we remind the reader of the aim of our paper. We are interested to assess whether there is complementarity or substitutability between parental inputs in middle childhood and child socio-emotional skills in early childhood (at age 6) in producing socio-emotional skills in middle childhood (at age 11). To test for this, we allow the productivity of each parental input to change across quartiles of socio-emotional skills in early childhood.²⁷ In addition and as an extension to previous economic papers on child skill formation, we also allow the productivity of each parental input to vary between high

²⁶To be more precise, $\theta_{i,1,2}^N$ and $\theta_{i,2,2}^N$ are factors derived from the measurement models described in Section 3.2.

²⁷We chose quartiles because finer partitions of the distribution would lead to imprecise estimates.

and low levels of the parental input as defined by levels above and below the median.²⁸

This high level of flexibility of our model is motivated by the psychological literature. Low levels of parental inputs can cause a stressful home environment that, when interacted with a child predisposition to psychological problems, can lead to the development of behavioural disorders as predicted by the psychological hypothesis *diathesis-stress*.²⁹ We therefore expect a child with low socio-emotional skills in early childhood, which can be an indicator of a predisposition to psychological problems, to be more vulnerable to a change in a parental input when this input is low and causing a more stressful home environment. Conversely, we expect a child with high socio-emotional skills in early childhood to be more resilient to a change in a parental input which is low. This would imply substitutability rather than complementarity for low levels of a parental input, as exemplified in Figure 1 by the steeper slope of the dotted line in Panel A with respect to the dotted line in Panel B for levels of inputs that are low.

Why might we expect complementarity for high levels of the inputs? Firstly, complementarity would be consistent with the early years of childhood being sensitive or critical periods for socio-emotional skill development, for high level of inputs (Cunha and Heckman 2008).³⁰ Secondly, high levels of parental inputs characterise an enriched home environment that, according to the psychological hypothesis *vantage sensitivity*, can represent an environment more beneficial for children with high rather than low level of skills in early childhood (see Pluess 2015). Consequently, a child with high socio-emotional skills in early childhood, which can be an indicator of vantage sensitivity, may be more sensitive to an increase in a parental input when this input is already high and enriched. This would imply complementarity for high level of parental inputs, as exemplified in Figure 1 by the flatter slope of the solid line in Panel A with respect to the solid line in Panel B for high level of inputs.

Table 4 reports our benchmark results separately for boys and girls in Panel A and B respectively. Because we use child fixed effect estimation we identify and report only the differential rather than the absolute productivity of each of the parental inputs for children at the second (column 1), third (column 2) and fourth (column 3) quartiles relative to the

²⁸When allowing the productivity of each parental input to vary across more fine partition of the distribution (e.g. deciles), we get estimates that, although imprecise, suggest that changes in the complementarity/substitutability pattern of parental inputs occurs only when moving from below to above the median.

²⁹This hypothesis is also known as disadvantage sensitivity because it implies a vulnerability to a negative home environment (see for a definition of such sensitivity hypothesis Monroe and Simons 1991; Belsky and Pluess 2009).

³⁰During sensitive periods parental investments into child development have relatively high returns and the critical periods of investments are the vital period during childhood for investments to occur.

first quartile of the socio-emotional skills in early childhood³¹ and separately for levels of the parental input below and above the median, which are reported in the first and second row of each sub-panel respectively. The socio-emotional skills at 6 and 11 and all parental inputs are standardized to have mean zero and variance one.³² A negative (positive) differential productivity at higher quartiles with respect to the first quartile indicates the presence of substitutability (complementarity). At the bottom of each of the two panels we report the self-productivity or persistence parameter for socio-emotional skills; this is computed keeping all other inputs at the mean and it represents the effect of one standard deviation increase of socio-emotional skills at 6 on socio-emotional skills at 11, also expressed in standard deviations.

Looking at boys and focusing on sensitive parenting style (see Table 4, Panel A, first sub-panel), we find substitutability for levels of the sensitive parenting style below the median and complementarity for levels above the median; these results are consistent with the implications of the diathesis-stress and vantage sensitivity hypotheses represented in Figure 1 by the dotted lines below the median and solid lines above the median.

In the first row of Panel A of Table 4, we find that an increase in sensitive parenting by one standard deviation for harsh parents (parents with a sensitive parenting style below the median) raises skills at age 11 by 30.0% and 27.5% of a standard deviation less for children in the third and fourth quartile of socio-emotional skills at age 6 compared to those in the first quartile. This implies that boys with low socio-emotional skills at age 6 (in the bottom two quartiles) are more vulnerable to a decrease in sensitive parenting compared to higher skilled children (in the third or fourth quartile) who are relatively resilient. These results are consistent with the diathesis-stress hypothesis, as low levels of sensitive parenting, i.e. harsh parental discipline such as smacking or shouting at the child, can increase the child's stress and especially affect children who have low socio-emotional skills.

On the other hand, an increase in parenting style for parents initially adopting a more sensitive style of parenting (sensitive parenting style above the median) would not indicate a movement away from a stressful environment but rather an even more enriched environment. The vantage sensitivity hypothesis would then suggest that children with relatively high socio-emotional skills at age 6 are more sensitive to such improvements in sensitive parenting. This is confirmed by our results which indicate that an improvement in sensitive parenting statistically significantly increases socio-emotional skills at age 11 by around 50.4% and 45.9%

³¹The quartiles are defined separately for boys and girls.

³²The only exceptions are dummy variable inputs which are not standardized and family income and mother's education which are demeaned.

of a standard deviation more for boys in the third and fourth quartile of the distribution at age 6 compared to those in the first quartile (see the second row of Panel A in Table 4).

For boys, we find evidence for neither complementarity nor substitutability in any other of the inputs, which imply that those inputs do not have a heterogeneous effect by socio-emotional skills. For brevity, Table 4 reports the coefficients only on inputs for which we found either substitutability or complementarity, for boys and/or for girls. Results for the remaining main inputs are reported in Appendix Table A.9, whereas results for the remaining control variables are reported in Appendix Table A.10.³³ Finally, looking at the self-productivity parameter (see the fourth sub-panel of Panel A in Table 4), we find that a standard deviation increase in socio-emotional skills at age 6 raises socio-emotional skills at age 11 by 47.0% of a standard deviation.

The pattern in the productivity of sensitive parenting style for girls is similar to the one observed for boys (see first sub-panel of Panel B in Table 4). We find substitutability between sensitive parenting style and early socio-emotional skills for levels of the parental input below the median and complementarity for levels of the input above the median, confirming again the implications of the psychological hypotheses represented in Figure 1 by the dotted lines below the median and solid line above the median. Notice that the productivity of sensitive parenting style at the first quartile is not statistically significantly different from the productivity at the second and third quartiles; but it is statistically significantly different than the productivity at the top quartile for both low and high levels of sensitive parenting. For girls, we also find a pattern in the productivity of time investment which is very similar to the one observed for sensitive parenting (see third sub-panel of Panel B in Table 4). This seems to suggest that levels of parental time investments below the median may cause a more stressful home environment for girls. Moreover, this result is also consistent with the evidence provided in Baker and Milligan (2016) finding that parents tend to invest more time in girls compared to boys.

Finally, for girls, we find even stronger evidence of substitutability between maternal mental health and early socio-emotional skills for levels of the maternal mental health below the median (see second sub-panel of Panel B in Table 4), again supporting the diathesis-stress hypothesis. Girls with early socio-emotional skills in the bottom quartile are more vulnerable to a decrease in mothers' mental health, as their socio-emotional skills at age 11

³³In a robustness check, available upon request to the authors, we also include school inputs measured by years of experience of the school teachers, which is a variable available only for a smaller sub-sample. The results show that this school input does not have a statistically significant differential effect across the socio-emotional skills at age 6.

decrease by 18.3-20.9% of a standard deviation more compared to girls with early skills in the top three quartiles. On the contrary, if maternal mental health is above the median, an increase of the maternal input is more beneficial for girls in the second or third quartile of the socio-emotional skills distribution at 6, compared to those in the bottom quartile who are relatively resistant to this improvement. That the differential effect of mothers' mental health at the bottom quartile of the child socio-emotional skills is statistically significant only for girls and not for boys is consistent with previous empirical studies which have found the intergenerational transmission of mental health to be stronger from mothers to girls than from mothers to boys. E.g., [Powdthavee and Vignoles \(2008\)](#) find a statistically insignificant effect of maternal negative emotional stress on boys, but a strong effect on girls, when the children are in adolescence.³⁴

Ultimately, our findings suggest that policy can help children who are lagging behind in terms of socio-emotional skills with interventions in middle childhood aiming at improving specific parental inputs but only if these interventions are targeted to parents with low levels of these inputs. To provide some information on the size of the potential effect of these interventions, we adopt a "back on the envelope" computation. This makes use of our estimated coefficients in Table 4 to predict the reduction in the socio-emotional skills gap at 11 caused by an increase in each of the inputs. The predicted reduction is computed comparing the predicted gap in socio-emotional skills at 11 with and without assuming an increase in a specific parental input. In Table 5 we show the predicted reduction in the gap in socio-emotional skills at 11 caused by an increase of 1 standard deviation in the parental input (reported at the start of the row), for parents with a level of the input below the median and holding all other inputs are at their mean. The gaps are the differences in socio-emotional skills at 11 between a child who was at the second (column 2), third (column 3) and fourth quartile (column 4) of the socio-emotional skills at 6 with respect to a child at the bottom quartile.³⁵

Looking at results for boys in Table 5, we find large reductions in the gaps when considering an increase in sensitive parenting which range between 21% and 34%, but much more modest reductions when considering equivalent increases in the mother's mental health or in the parental time. Results for girls suggest that an increase by one standard deviation in sensitive parenting style, mother's mental health and parental time investments leads to substantial reductions in the gaps in socio-emotional skills at 11 for girls. This is especially

³⁴Results for the remaining main inputs are reported in Appendix Table A.9, whereas results for the remaining control variables are reported in Appendix Table A.10.

³⁵When considering a child in the q -th quartile of the socio-emotional skills at 6, we assume that his/her level of socio-emotional skills is equal to the average level observed for children in the q -th quartile.

true when increasing the mother’s mental health which could close up to 52% of the gap. In conclusion, results of this “back on the envelope” computation provides further evidence that interventions aiming at increasing certain parental inputs, when they are low, could be cost-effective to help children to catch-up their socio-emotional skills.

6 Validity of Model Assumptions

In this section we provide supportive evidence for the three main identification assumptions of our model: (i) invariance of the production model across the two trait-specific socio-emotional skills; (ii) random variation in the child predisposition to either internalising or externalising problems in early life; (iii) absence of omitted inputs which are correlated with the included inputs and whose productivity varies across the distribution of early socio-emotional skills.

First, the latent traits for externalising and internalising are assumed to follow the same dynamic production model. In practice this means assuming invariance of the effect of inputs, shocks and early childhood skills across externalising and internalising traits.

The assumption that the effect of inputs and shocks is the same for externalising as internalising traits may be violated if for example, the response to a particular shock leads to children internalising over externalising, or vice versa. However, there is evidence from the economics and psychology literatures of a similar reaction of externalising and internalising traits in response to shocks to income (Akee et al., 2018), divorce (Lansford et al., 2006) and parent maltreatment (Anthonysamy and Zimmer-Gembeck, 2007; Maschi et al., 2008). To give some more detail, Akee et al. (2018) analyse the effect of an unconditional cash transfer to participants of the Great Smoky Mountains study. The authors find very similar effects between emotional and behavioural problems, corresponding to our measures of internalising and externalising behaviour. Similarly Lansford et al. (2006) find very similar patterns when examining how parental divorce affects the average development trajectory of externalising compared to internalising traits. Finally, Anthonysamy and Zimmer-Gembeck (2007) has documented that externalising and internalising problems develop in a similar manner following episodes of child abuse.

The assumption of invariance of the persistence is not confirmed by previous empirical studies, which find that externalising behaviours tend to decrease across child age, whereas the internalising behaviours tend to increase (see e.g. Bornstein et al. 2010; Gilliom and

Shaw 2004). For this reason, we also consider the estimation of a model that allows the self-productivity coefficient to change between internalising and externalising, with results reported in Appendix Table A.11. We do not reject the assumption of invariance of the persistence between the two trait-specific socio-emotional skills at 5% level, neither for boys nor for girls. In addition, the differential productivity coefficients of parental inputs remain similar and our conclusions relating to the interactions between inputs and lagged socio-emotional skills remain the same.

Secondly, we assume that the variation in the deviation of a trait-specific socio-emotional skill from the mean across the two traits, which we use when we consider the child fixed effect estimation, picks up a random predisposition to externalising or internalising behaviours. To understand if this is the case, we consider all potential sources of variation in the differences between externalising and internalising behaviours at age 6. These differences may be due to an innate random predisposition, as we assume, or to three other potential sources which we now discuss in turn.

It could be that the within child deviation in lagged externalising and internalising traits comes from differences in trait-specific inputs. The only trait-specific inputs that could explain the trait-specific socio-emotional skills at age 6 and 11 are the trait-specific socio-emotional skills of the parents, which can be transmitted to children. In a sensitivity analysis we estimate our model including both mothers' and fathers' trait-specific socio-emotional skills to control for such intergenerational transmission (see Panel A and B in Table A.12 for boys and girls respectively).³⁶ The conclusions from our benchmark results are unchanged once we control for parent trait-specific socio-emotional skills. It is not expected that other individuals who interact with the child, for example teachers, would make trait-specific investments in children. We conclude that the variation in child's trait-specific socio-emotional skills are not caused by other trait specific investments.

Next, the differences between the measures of internalising and externalising behaviours can be caused by measurement errors in the mother reports of child socio-emotional skills. However, in our main specification we control for this by estimating a latent factor, and furthermore Section 7 below shows that our results are robust to measures of child behaviour using teachers reports.

Moreover, there may be differences in the effects of past investments and inputs between externalising and internalising behaviours. Nevertheless, as discussed above, this is incon-

³⁶The subsample with details on both father and mother trait-specific skills is smaller. Table A.13 shows descriptive statistics for such subsample.

sistent with the evidence which shows that events including family income shocks (Akee et al., 2018), divorce (Lansford et al., 2006) and parent maltreatment (Anthonyamy and Zimmer-Gembeck, 2007; Maschi et al., 2008) induce a similar reaction in externalising and internalising behaviours. Furthermore, even in presence of past investments and inputs with effects that vary between externalising and internalising behaviours at age 6, their omission would not cause any endogeneity bias for the estimation of the production model at age 11 as long as they do not have a direct effect on socio-emotional skills at 11 which varies between the two traits.

Therefore, given that we can rule out all other potential sources of endogenous variation in the differences between externalising and internalising behaviours at age 6, we conclude that the source comes primarily from a random innate predisposition to internalising and externalising behaviours.

Finally, our third identification assumption is that there are no omitted inputs in our model whose returns change across levels of the socio-emotional skills at 6. This assumption is needed to guarantee the exogeneity of the parental inputs in our estimation model. Because we adopt a child fixed effect estimation, we are able to control for any omitted input which is invariant across the two trait-specific socio-emotional skills. Nevertheless, omitted inputs whose effects differ across the distribution of early socio-emotional skills can still bias our estimation if they are correlated with our parental inputs of interest. As explained in Section 3.1, our model includes a much larger set of inputs than just the parental inputs of interests. Our results in Appendix Table A.10 show there is no evidence of differential returns to these inputs across early socio-emotional skills even when considering events which could cause a shock in the child’s life and therefore some behavioural reactions, such as the birth of another child (fertility), having a mother with poor health or a father with longstanding illness. In addition, it may be that parental inputs measured before the child was aged 6 can cause an omitted variable bias. Therefore in Tables A.14 and A.15 we control additionally for the three main inputs of sensitive parenting, time investments and mothers’ mental health, measured when the child was aged 3. Our results are robust to this addition, indicating there are no concerns about the endogeneity of the parental inputs.

7 Sensitivity Analysis

To ensure that there are no other endogeneity issues threatening our identification strategy, we run several robustness checks where we (i) use mother and teacher reports to correct

for measurement error; (ii) test for differences in self-productivity across quartiles of socio-emotional skills at age 6; (iii) anchor socio-emotional skills to a real-life measure.

A first sensitivity analysis controls for the measurement error in the child socio-emotional skills reported by the mother. The mother reports could be systematically biased and correlated with mothers mental health as found by [Del Bono et al. \(2017\)](#). Controlling for the measurement error in the child socio-emotional skills, by using both mother’s and teacher’s reports to derive the latent factors for internalising and externalising behaviours, does not seem cause any large change in the estimated coefficients (see Table [A.16](#)).³⁷ However, the estimates are not identical and are less precisely estimated, most likely due to the smaller sample size. We conclude that there is no systematic empirical bias for our benchmark results. The lack of measurement error bias in our estimates is probably explained by the fact that we use latent factors for the two socio-emotional skills and that the child fixed effect estimation exploits variation in socio-emotional skills within child and across traits and therefore corrects for any potential measurement error which is shared across different measures of socio-emotional skills reported by the mother.

Our second sensitivity analysis relaxes the assumption of linearity in the socio-emotional skills at 6 imposed by our benchmark model (6) by allowing the self-productivity coefficient ρ to vary across different quartiles of the socio-emotional skills at 6. Results of this specification are shown Table [A.17](#) in the Appendix for boys and girls, in Panel A and B respectively. We do not reject the hypothesis of equality of the self-productivity coefficient across quartile at 5% level of significance, therefore the linearity assumption in our benchmark model is not rejected.

Finally, the lack of a natural metric of our socio-emotional skills implies that our results might be sensitive to arbitrary scaling of socio-emotional skills. This is especially an issue when regressing an arbitrarily scaled outcomes on arbitrarily scaled lagged outcomes (see [Cunha and Heckman 2008](#); [Cunha et al. 2010](#); [Bond and Lang 2013](#)).³⁸ We follow [Cunha et al. \(2010\)](#) and re-scale socio-emotional skills by anchoring them to a measure with a well defined scale, which in our case is smoking behaviour at age 14. Using the latest wave of the MCS, we construct an indicator of whether the child has ever smoked at age 14, taking value 0 if the child has smoked and 1 if the child has never smoked. We regress the

³⁷We use a latent factor model to combine measures of socio-emotional skills using the repeated measures reported by the mother at age 5 and 7 and the measures reported by the teacher at age 7.

³⁸Notice that arbitrary rescaling is less of an issue when considering the different quartiles of the socio-emotional skills at age 6 given that any monotonic transformation of the socio-emotional skills would yield the same values for each of these quartiles.

smoking dummy separately on internalising and externalising behaviour at ages 6 and 11 using linear spline regression. Then we import the estimated coefficients into our estimation sample and perform our analysis on socio-emotional skills transformed into the probability of non-smoking.

Table A.18 in the Appendix displays the results for boys and girls, in Panel A and B respectively. Looking at the degree of self-productivity, we find an increase in socio-emotional skills at age 6 that corresponds to 1 percent increase in the probability of non-smoking raises the socio-emotional skills at age 11 of an amount that corresponds to an increase in the probability of non-smoking of 0.44 percentage points for boys and of 0.30 percentage points for girls. These results confirm that the higher self-productivity observed among boys compared to girls in our benchmark results is not affected by the scale of the observed measures. The anchoring exercise confirms all our benchmark results including the evidence on substitutability and complementarity between the parental inputs and socio-emotional skills at age 6 in producing socio-emotional skills at age 11.

8 Conclusions

Existing empirical evidence suggest that inequalities in socio-emotional and cognitive skills are present before children start schooling and persist across their childhood. However, socio-emotional skills tend to be malleable much further into childhood than cognitive skills; this suggests that these early life differences can be attenuated through interventions in middle childhood, e.g. introducing public policies that support parents to improve their parenting behaviour.

To assess the potential usefulness of such interventions in middle childhood and to identify which types of parental inputs should be targeted, we evaluate the degree of complementarity and/or substitutability between different types of parental inputs and socio-emotional skills at age 6 in producing socio-emotional skills at age 11. In particular, we test whether there is heterogeneity across quartile of socio-emotional skills at age 6 in the productivity of the following six parental inputs: sensitive parenting style, routines parenting style, time investment, family income, mother’s cognitive skills and mother’s mental health.

We find evidence of substitutability between socio-emotional skills at 6 and parental inputs that help reducing child’s stress, for both boys and girls. Children with low socio-emotional skills early in life are more vulnerable to a stressful home environment, such as

harsh parenting, poor mothers' mental health and low parental time investment. An improvement of these inputs when the initial values these parental inputs is below the median is more beneficial for children with low rather than high socio-emotional skills in early childhood. In line with existing evidence in the child development literature our findings suggest that the early years are crucial for the development of socio-emotional skills, as children with low skills at age 6 are set on a different trajectory of skill development up to age 11 compared to those with high early skills.

Coming back to our initial question of whether interventions in middle childhood can be effective to narrow the gaps in socio-emotional skills, the short answer is yes. Nevertheless, to be cost-effective such interventions should be aimed at raising parental inputs that reduce stressful environment experiences for the child and should be targeted exclusively to parents with low level of such parental inputs. Specifically, a policy aiming at reducing harsh (increasing sensitive) parenting, improving mother's mental health and increasing parental time investment targeted to parents whose level of these inputs is below the median and therefore indicative of a stressful home environment, would be effective in shrinking gaps in socio-emotional skills of children in middle childhood.

References

- Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: a factor-analytic study. *Psychological Monographs: general and applied*, 80(7):1.
- Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., and Althoff, R. R. (2016). Internalizing/externalizing problems: Review and recommendations for clinical and research applications. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(8):647–656.
- Agostinelli, F. and Wiswall, M. (2016). Estimating the technology of children's skill formation. *National Bureau of Economic Research, (NBER) Working Paper No. 22442*.
- Akee, R., Copeland, W., Costello, E. J., and Simeonova, E. (2018). How does household income affect child personality traits and behaviors? *American Economic Review*, 108(3):775–827.
- Anthonyamy, A. and Zimmer-Gembeck, M. J. (2007). Peer status and behaviors of maltreated children and their classmates in the early years of school. *Child Abuse & Neglect*, 31(9):971–991.

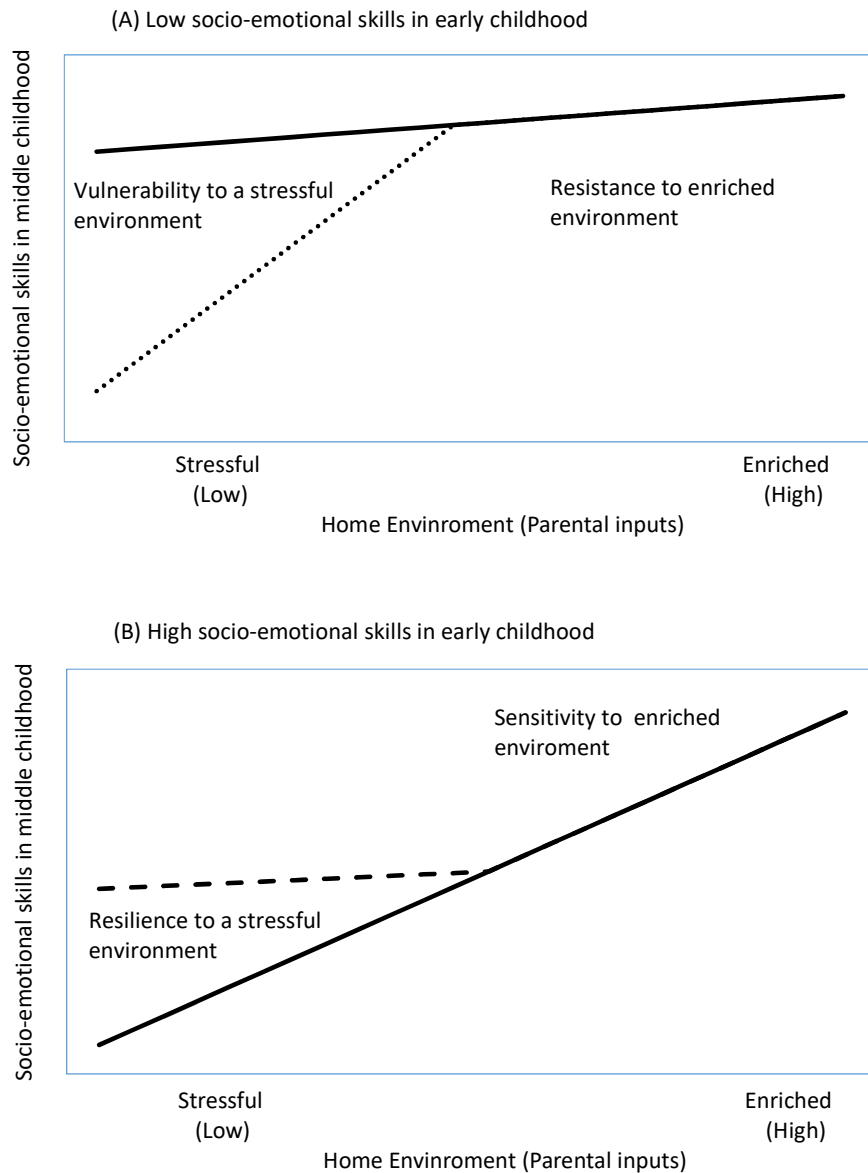
- Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2015). Estimating the production function for human capital: Results from a randomized control trial in Colombia. *National Bureau of Economic Research, (NBER) Working Paper No. 20965*.
- Attanasio, O., Meghir, C., Nix, E., and Salvati, F. (2017). Human capital growth and poverty: Evidence from Ethiopia and Peru. *Review of Economic Dynamics*, 25:234–259.
- Baker, M. and Milligan, K. (2016). Boy-girl differences in parental time investments: Evidence from three countries. *Journal of Human Capital*, 10(4):399–441.
- Bandiera, O., Larcinese, V., and Rasul, I. (2010). Heterogeneous class size effects: New evidence from a panel of university students. *The Economic Journal*, 120(549):1365–1398.
- Belsky, J. and Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological bulletin*, 135(6):885.
- Bertrand, M. and Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.
- Bond, T. N. and Lang, K. (2013). The evolution of the black-white test score gap in grades k–3: The fragility of results. *Review of Economics and Statistics*, 95(5):1468–1479.
- Bornstein, M. H., Hahn, C.-S., and Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and psychopathology*, 22(4):717–735.
- Carneiro, P. and Heckman, J. (2003). Human capital policy. *National Bureau of Economic Research, (NBER) Working Paper No. 9495*.
- Carneiro, P. M., Garcia, I. L., Salvanes, K. G., and Tominey, E. (2015). Intergenerational mobility and the timing of parental income. *CEMMAP working paper CWP66/15*.
- Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3):655–681.
- Cobb-Clark, D. A., Salamanca, N., and Zhu, A. (2016). Parenting style as an investment in human development. *Discussion paper series, IZA DP No. 9686*.
- Cunha, F. and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4):738–782.

- Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.
- Dahl, G. B. and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5):1927–56.
- Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3):528–554.
- Del Boca, D., Flinn, C., and Wiswall, M. (2013). Household choices and child development. *Review of Economic Studies*, 81(1):137–185.
- Del Boca, D., Monfardini, C., and Nicoletti, C. (2017). Parental and child time investments and the cognitive development of adolescents. *Journal of Labor Economics*, 35(2):565–608.
- Del Bono, E., Francesconi, M., Kelly, Y., and Sacker, A. (2016). Early maternal time investment and early child outcomes. *The Economic Journal*, 126(596).
- Del Bono, E., Kinsler, J., and Pavan, R. (2017). Estimating the technology of skill formation and maternal well-being. *mimeo*.
- Del Giudice, M. (2018). Middle childhood: An evolutionary-developmental synthesis. In *Handbook of life course health development*, pages 95–107. Springer, Cham.
- Dex, S. and Joshi, H. (2005). *Children of the 21st century: from birth to nine months*. Policy Press.
- Doepke, M. and Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371.
- Elliott, C. D., Smith, P., and McCulloch, K. (1997). British ability scale, second edition. administration and scoring manual. Technical report, London:NFER-Nelson.
- Fiorini, M. and Keane, M. P. (2014). How the allocation of children’s time affects cognitive and noncognitive development. *Journal of Labor Economics*, 32(4):787–836.
- García, J. L. and Gallegos, S. (2017). Dynamic complementarity or substitutability? parental investment and childcare in the production of early human capital. *Available at SSRN: <https://ssrn.com/abstract=2910167> or <http://dx.doi.org/10.2139/ssrn.2910167>*.
- Gilliom, M. and Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and psychopathology*, 16(2):313–333.

- Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, 38(5):581–586.
- Hansen, K., Johnson, J., Joshi, H., et al. (2012). Millennium Cohort Study, first, second, third and fourth surveys: a guide to the datasets. *London: Centre for Longitudinal Studies, Institute of Education*.
- Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6(1):689–733.
- Heckman, J. J. and Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the ged testing program. *The American Economic Review*, 91(2):145–149.
- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.
- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. *National Bureau of Economic Research, (NBER) Working Paper No. 20749*.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L., Walters, E. E., and Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological medicine*, 32(6):959–976.
- Lansford, J. E., Malone, P. S., Castellino, D. R., Dodge, K. A., Pettit, G. S., and Bates, J. E. (2006). Trajectories of internalizing, externalizing, and grades for children who have and have not experienced their parents’ divorce or separation. *Journal of family Psychology*, 20(2):292.
- Lavy, V., Silva, O., and Weinhardt, F. (2012). The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2):367–414.
- Liu, H., Mroz, T. A., and Van der Klaauw, W. (2010). Maternal employment, migration, and child development. *Journal of Econometrics*, 156(1):212–228.
- Maschi, T., Morgen, K., Bradley, C., and Hatcher, S. S. (2008). Exploring gender differences on internalizing and externalizing behavior among maltreated youth: Implications for social work action. *Child and Adolescent Social Work Journal*, 25(6):531–547.

- Monroe, S. M. and Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin*, 110(3):406.
- Moroni, G. (2018). Explaining divorce gap in cognitive and non-cognitive skills of children. *University of York, Discussion Papers in Economics, No 18/16*.
- Mueller, S. (2013). Teacher experience and the class size effect—experimental evidence. *Journal of Public Economics*, 98:44–52.
- Nicoletti, C. and Rabe, B. (2014). School inputs and skills: Complementarity and self-productivity. *IZA Discussion Paper No. 8693*.
- Nicoletti, C. and Rabe, B. (2017). The effect of school spending on student achievement: addressing biases in value-added models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(2):487–515.
- Pluess, M. (2015). Individual differences in environmental sensitivity. *Child Development Perspectives*, 9(3):138–143.
- Powdthavee, N. and Vignoles, A. (2008). Mental health of parents and life satisfaction of children: A within-family analysis of intergenerational transmission of well-being. *Social Indicators Research*, 88(3):397–422.
- Straus, M. A. and Hamby, S. L. (1997). Measuring physical & psychological maltreatment of children with the conflict tactics scales. In G. K. Kantor & J.L.Jasinski (Eds.) *Out of the darkness: Contemporary research perspectives on family violence*.

Figure 1: Effect of home environment on child socio-emotional skills in middle childhood



Notes: Panels (A) and (B) plot the regression line of socio-emotional skills in middle childhood on home environment (parental inputs) for low and high skilled children in early childhood. The difference in the slopes of the two solid lines in Panel (A) and (B) can be explained by the presence of complementarity between inputs in middle childhood and skills in early childhood as suggested by previous economic studies.

Table 1: SDQ Questionnaire. Details on items for each scale.

Scale	Items within scale
1) Emotional symptoms	<ul style="list-style-type: none"> a. Complaints of headaches/stomach aches/sickness b. Often seems worried c. Often unhappy d. Nervous or clingy in new situations e. Many fears, easily scared.
2) Peer Problems	<ul style="list-style-type: none"> a. Tends to play alone b. Has at least one good friend* c. Generally liked by other children* d. Picked on or bullied by other children e. Gets on better with adults.
3) Conduct problems	<ul style="list-style-type: none"> a. Often has temper tantrums b. Generally obedient* c. Fights with or bullies other children d. Can be spiteful to others e. Often argumentative with adults.
4) Hyperactivity	<ul style="list-style-type: none"> a. Restless, overactive, cannot stay still for long b. Constantly fidgeting c. Easily distracted d. Can stop and think before acting* e. Sees tasks through to the end*.

Notes: For ease of interpretation each scale is reverse coded, with higher values associated with higher levels of socio-emotional skills (or lower levels of socio-emotional problems). * denotes items that are reversed when computing the total score for each type of behavioural problem (scale). All of the 25 items are related with a statement on the child and the parent has to indicate whether the statement is 'true', 'somewhat true' or 'not true'. Each of the 4 scales can be used alone or together to create: (i) total difficulties score when 1-4 are taken together ; (ii) an internalising behavioural score combining 1 and 2 ; (iii) an externalising behavioural score combining 3 and 4.

Table 2: Descriptive statistics of child’s skills, Raw Measures

	Boys		Girls		Test diff
	mean	sd	mean	sd	p-value
Socio-emotional skills age 5					
<i>Internalising behaviour</i>					
Emotional Symptoms	8.813	1.474	8.765	1.434	0.181
Peer Problems	8.976	1.369	9.128	1.226	0.000
<i>Externalising behaviour</i>					
Conduct Problems	8.605	1.401	8.827	1.295	0.000
Hyperactivity Problems	6.718	2.297	7.406	2.098	0.000
Socio-emotional skills age 7					
<i>Internalising behaviour</i>					
Emotional Symptoms	8.699	1.621	8.631	1.575	0.090
Peer Problems	8.973	1.444	9.100	1.288	0.000
<i>Externalising behaviour</i>					
Conduct Problems	8.704	1.440	8.983	1.286	0.000
Hyperactivity Problems	6.539	2.437	7.445	2.243	0.000
Socio-emotional skills age 11					
<i>Internalising behaviour</i>					
Emotional Symptoms	8.424	1.857	8.234	1.910	0.000
Peer Problems	8.810	1.604	8.938	1.508	0.001
<i>Externalising behaviour</i>					
Conduct Problems	8.693	1.514	8.937	1.349	0.000
Hyperactivity Problems	6.759	2.457	7.698	2.134	0.000
Cognitive skills age 5					
BAS:Picture similarities	82.669	10.995	83.906	10.886	0.000
BAS:Reading abilities	111.198	14.769	111.750	13.986	0.123
BAS:Pattern construction	89.204	18.492	91.604	16.848	0.000
Cognitive skills age 7					
NFER: Maths test	10.119	2.706	9.977	2.511	0.029
BAS:Reading abilities	108.986	30.615	112.721	26.530	0.000
BAS:Pattern construction	118.040	16.432	119.474	15.167	0.000
Observations	3217		3235		6452

Notes: Last column reports the p-value of the test for the equality of means between boys and girls. Data: UK Millennium Cohort Study, Main Sample.

Table 3: Descriptive statistics of all the inputs in stage 2, by gender

	Boys		Girls		Test diff
	mean	sd	mean	sd	p-value
Main inputs					
<i>Parental investment</i>					
Sensitive parenting style	-0.128	0.996	0.126	0.989	0.000
Routines parenting style	0.007	0.981	-0.006	1.016	0.600
Time investment	-0.042	1.010	0.041	0.988	0.001
Family income	496.994	242.311	496.796	245.115	0.974
<i>Parental skills</i>					
Mother's education	18.142	2.605	18.138	2.563	0.944
Mother's mental health	-0.003	0.989	-0.000	1.014	0.932
<i>Mother's general health</i>					
Excellent	0.257	0.437	0.250	0.433	0.500
Good	0.649	0.477	0.655	0.476	0.633
Fair	0.074	0.262	0.075	0.264	0.862
Poor	0.020	0.139	0.020	0.140	0.883
Health Capital					
<i>Child longstanding illness</i>					
Yes	0.201	0.401	0.147	0.354	0.000
<i>Father longstanding illness</i>					
Yes	0.228	0.420	0.236	0.425	0.447
Neighbourhood Capital					
Income decile area	5.958	2.856	5.891	2.895	0.351
Social Capital					
<i>Any family and/or friends nearby</i>					
Yes friends	0.234	0.423	0.226	0.418	0.475
Yes family	0.041	0.198	0.044	0.204	0.610
Yes both	0.682	0.466	0.697	0.460	0.191
No	0.043	0.203	0.033	0.179	0.034
Other Family decisions					
Non-parental Child care (Hours)	4.573	8.489	4.813	8.701	0.263
Hours Club	0.461	0.992	0.443	0.963	0.459
Working hours per week	17.439	14.027	17.477	14.664	0.916
Fertility	0.134	0.340	0.142	0.349	0.321
Observations	3217		3235		6452

Notes: Sensitive parenting style, Routines parenting style, Time investment, Mother's mental health and Child's cognitive skills are obtained from factor models. They are standardized to have mean 0 and standard deviation 1. Family income, that is the OECD equivalised weekly family income, is adjusted at 2012 prices. Last column reports the p-value of the test for the equality of means between boys and girls. Data: UK Millennium Cohort Study, Main Sample.

Table 4: Production model of socio-emotional skills (between age 6 and 11), by gender.

	Differential effect of the input at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
relative to the 1st Quartile of socio-emotional skills			
Panel A: Boys			
Sensitive parenting style \leq median	-0.128 (0.089)	-0.300*** (0.091)	-0.275*** (0.097)
Sensitive parenting style $>$ median	0.247 (0.170)	0.504*** (0.168)	0.459*** (0.171)
Mother's mental health \leq median	-0.067 (0.067)	-0.001 (0.061)	-0.099 (0.090)
Mother's mental health $>$ median	0.032 (0.198)	-0.141 (0.194)	-0.006 (0.217)
Time investment \leq median	0.059 (0.084)	-0.010 (0.083)	0.068 (0.087)
Time investment $>$ median	0.008 (0.151)	0.022 (0.154)	-0.095 (0.162)
Socio-emotional skills (age 6)		0.470*** (0.045)	
σ^2 (variance of fixed effect)		0.467	
Observations		6434	
Panel B: Girls			
Sensitive parenting style \leq median	-0.022 (0.096)	-0.121 (0.108)	-0.245*** (0.110)
Sensitive parenting style $>$ median	0.060 (0.162)	0.185 (0.169)	0.353*** (0.170)
Mother's mental health \leq median	-0.209*** (0.057)	-0.198*** (0.069)	-0.183*** (0.074)
Mother's mental health $>$ median	0.589*** (0.185)	0.596*** (0.196)	0.253 (0.200)
Time investment \leq median	-0.111 (0.090)	-0.093 (0.093)	-0.180* (0.096)
Time investment $>$ median	0.235 (0.162)	0.112 (0.160)	0.345** (0.166)
Socio-emotional skills (age 6)		0.307*** (0.047)	
σ^2 (variance of fixed effect)		0.467	
Observations		6470	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. The results for the remaining variables are reported in Table A.9 and A.10 in the Appendix. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, Main Sample.

Table 5: Predicted reduction in the gap in socio-emotional skills at 11 for an increase in inputs below the median of 1 standard deviation.

	Percentage reduction in the gap between children at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
	and the 1st Quartile of socio-emotional skills at age 6		
Panel A:Boys			
Sensitive parenting style	20.898 (14.914)	33.737 (11.049)	24.245 (9.020)
Mother’s mental health	10.953 (10.960)	0.093 (6.813)	8.736 (7.994)
Time investment	-9.669 (13.733)	1.154 (9.342)	-5.968 (7.718)
Panel B:Girls			
Sensitive parenting style	5.493 (24.073)	20.668 (19.016)	33.224 (16.155)
Mother’s mental health	52.279 (17.023)	33.783 (13.171)	24.814 (10.900)
Time investment	27.842 (23.086)	15.939 (15.994)	24.337 (13.414)

Notes: Percentage reductions in gaps are computed as the difference between the predicted gap with and without an increase of one standard deviation in the input. The predictions are computed using our benchmark model estimated coefficients and keeping all other inputs at the mean. Standard errors in parenthesis.

A Appendix

Table A.1: Descriptive statistics of demographic characteristics in stage 2 of the main sample.

	mean	sd
Demographic characteristics		
Age in months	86.611	2.870
Number of Siblings	1.510	0.964
Mother age	37.076	5.304
<i>Mother's ethnicity</i>		
White	0.929	0.256
Mixed	0.005	0.068
Indian	0.019	0.137
Pakistani Bangladeshi	0.023	0.151
Black	0.014	0.116
Other	0.010	0.100
<i>Mother's occupational status</i>		
Managerial and professional	0.364	0.481
Intermediate	0.290	0.454
Routine and manual	0.316	0.465
Never worked	0.030	0.170
Observations	6452	

Notes: The demographic characteristics in stage 2 are measured around age 7 of the child. Data: UK Millennium Cohort Study, Main Sample.

Table A.2: Factor loadings of Internalising and Externalising behaviour - Boys

	(1) Factor Loadings	(2) Signal
Internalising behaviour		
Peer problems (age 5)	1.000	0.479
Peer problems (age 7)	1.062	0.487
Emotional symptoms (age 5)	0.894	0.332
Emotional symptoms (age 7)	1.059	0.384
Externalising behaviour		
Hyperactivity problems (age 5)	1.000	0.637
Hyperactivity problems (age 7)	1.084	0.665
Conduct problems (age 5)	0.489	0.408
Conduct problems (age 7)	0.513	0.428

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent. Data: UK Millennium cohort Study, Main Sample.

Table A.3: Factor loadings of Internalising and Externalising behaviour - Girls

	(1) Factor Loadings	(2) Signal
Internalising behaviour		
Peer problems (age 5)	1.000	0.393
Peer problems (age 7)	1.024	0.374
Emotional symptoms (age 5)	1.144	0.376
Emotional symptoms (age 7)	1.304	0.405
Externalising behaviour		
Hyperactivity problems (age 5)	1.000	0.607
Hyperactivity problems (age 7)	1.103	0.641
Conduct problems (age 5)	0.492	0.386
Conduct problems (age 7)	0.497	0.400

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent.
Data: UK Millennium cohort Study, Main Sample.

Table A.4: Factor loadings of child's cognitive skills - Boys

	(1) Factor Loadings	(2) Signal
BAS:Picture similarities (age 5)	1.000	0.220
BAS:Reading abilities (age 5)	1.420	0.245
BAS:Pattern construction (age 5)	2.396	0.445
NFER: Maths test (age 7)	0.360	0.470
BAS:Reading abilities (age 7)	3.398	0.327
BAS:Pattern construction (age 7)	2.099	0.433

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent.
Data: UK Millennium cohort Study, Main Sample.

Table A.5: Factor loadings of child's cognitive skills - Girls

	(1) Factor Loadings	(2) Signal
BAS:Picture similarities	1.000	0.210
BAS:Reading abilities	1.444	0.265
BAS:Pattern construction	2.124	0.395
NFER: Maths test	0.354	0.494
BAS:Reading abilities	2.929	0.302
BAS:Pattern construction	2.088	0.471

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent.
Data: UK Millennium cohort Study, Main Sample.

Table A.6: Factor loadings of parental investments

	(1) Factor Loadings	(2) Signal
Sensitive parenting style		
How often tells child off when naughty	1.000	0.502
How often smacks at child when naughty	0.473	0.175
How often shouts at child when naughty	0.974	0.421
How often sends child to bedroom/naughty chair	0.996	0.398
How often takes away treats from child when naughty	0.890	0.380
How often bribes at child when naughty	0.533	0.107
How often ignores at child when naughty	0.696	0.161
Routines parenting style		
Regular bedtime on term-time weekdays	1.000	0.289
What bedtime on term-time weekdays	1.317	0.254
Rules about timed watching TV	0.331	0.131
Rules about hours watching TV	0.370	0.083
Time investment		
Frequency draw and paint with child	1.000	0.372
Frequency tell stories to child	0.958	0.191
How often does musical activities with child	0.868	0.183
Frequency play active games with child	1.091	0.347
Frequency play indoor games with child	1.089	0.440
Frequency take child to park or playground	0.617	0.159

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent.
Data: UK Millennium cohort Study, Main Sample.

Table A.7: Factor loadings of mother's socio-emotional skills

	(1) Factor Loadings	(2) Signal
Mother's mental health		
How often felt depressed in last 30 days	1.000	0.629
How often felt hopeless in last 30 days	0.981	0.707
How often felt restless in last 30 days	0.830	0.368
How often felt everything an effort in last 30 days	1.040	0.473
How often felt worthless in last 30 days	0.954	0.626
How often felt nervous in last 30 days	0.751	0.312

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent.
Data: UK Millennium cohort Study, Main Sample.

Table A.8: Factor loadings of parents' trait specific socio-emotional skills

	(1) Factor Loadings	(2) Signal
Mother's internalising		
I get overwhelmed by emotions	1.000	0.444
I get caught up in my problems	0.962	0.478
I bottle up my feelings	-0.844	0.285
I am a very private person	-0.510	0.119
Mother's externalising		
I grumble about things	1.000	0.330
I get angry easily	1.109	0.420
I take offence easily	0.868	0.253
Father's internalising		
I get overwhelmed by emotions	1.000	0.254
I get caught up in my problems	1.360	0.456
I bottle up my feelings	-1.283	0.317
I am a very private person	-0.809	0.144
Father's externalising		
I grumble about things	1.000	0.293
I get angry easily	1.149	0.443
I take offence easily	0.833	0.250

Notes: Column (2) indicates for each of the measures the fraction of the variance explained by the latent. Data: UK Millennium Cohort Study, subsample with details on mother's and father's trait specific socio-emotional skills.

Table A.9: Production model of socio-emotional skills (between age 6 and 11), by gender. Remaining inputs effects.

	Differential effect of the input at the		
	(1) 2nd Quartile relative to the 1st	(2) 3rd Quartile Quartile of socio-emotional skills	(3) 4th Quartile
Panel A: Boys			
Routines parenting style \leq median	-0.081 (0.074)	-0.054 (0.075)	-0.097 (0.075)
Routines parenting style $>$ median	0.014 (0.162)	0.014 (0.166)	0.152 (0.164)
Family income \leq median	0.001 (0.047)	0.065 (0.049)	0.098* (0.052)
Family income $>$ median	0.013 (0.069)	-0.054 (0.070)	-0.099 (0.073)
Cognitive skills (age 6) \leq median	0.027 (0.081)	0.064 (0.088)	-0.023 (0.098)
Cognitive skills (age 6) $>$ median	0.071 (0.162)	-0.100 (0.161)	0.141 (0.165)
Mother's education \leq median	0.018 (0.050)	0.021 (0.050)	0.007 (0.051)
Mother's education $>$ median	0.003 (0.068)	-0.028 (0.066)	-0.019 (0.068)
Observations		6434	
Panel B: Girls			
Routines parenting style \leq median	0.113 (0.072)	0.000 (0.071)	0.028 (0.080)
Routines parenting style $>$ median	-0.310** (0.156)	0.049 (0.157)	-0.124 (0.163)
Family income \leq median	0.064 (0.049)	-0.037 (0.050)	-0.029 (0.051)
Family income $>$ median	-0.067 (0.072)	-0.007 (0.070)	0.005 (0.071)
Cognitive skills (age 6) \leq median	0.065 (0.100)	-0.042 (0.097)	-0.012 (0.101)
Cognitive skills (age 6) $>$ median	-0.083 (0.186)	0.089 (0.182)	0.083 (0.195)
Mother's education \leq median	-0.060 (0.053)	-0.009 (0.054)	0.029 (0.054)
Mother's education $>$ median	0.057 (0.074)	0.040 (0.074)	-0.018 (0.073)
Observations		6470	

Notes: The table shows the full set of interactions of the benchmark model shown in Table 4. Data: UK Millennium Cohort Study, Main Sample.

Table A.10: Production model of socio-emotional skills (between age 6 and 11), by gender. Remaining controls.

	Differential effect of the input at the		
	(1) 2nd Quartile	(2) 3rd Quartile	(3) 4th Quartile
	relative to the 1st Quartile of socio-emotional skills		
Panel A: Boys			
Mother General Health - Excellent	-0.363 (0.355)	0.251 (0.350)	-0.214 (0.341)
Mother General Health - Good	-0.298 (0.346)	0.280 (0.341)	-0.097 (0.336)
Mother General Health - Fair	-0.348 (0.393)	0.254 (0.391)	-0.334 (0.377)
Mother General Health - Poor	-0.155 (0.533)	0.356 (0.537)	-0.247 (0.518)
Child longstanding illness	-0.070 (0.134)	-0.072 (0.129)	-0.190 (0.125)
Father longstanding illness	-0.058 (0.125)	-0.152 (0.122)	0.113 (0.120)
Income decile area	-0.005 (0.019)	-0.030 (0.019)	-0.020 (0.019)
Family and/or friends nearby - Yes friends	-0.000 (0.278)	-0.153 (0.264)	-0.137 (0.244)
Family and/or friends nearby - Yes family	0.193 (0.328)	-0.118 (0.338)	0.046 (0.302)
Family and/or friends nearby - Yes both	0.044 (0.266)	-0.284 (0.251)	-0.081 (0.230)
Non-parental child care	0.002 (0.006)	0.006 (0.006)	0.006 (0.006)
Hours club	-0.055 (0.047)	-0.011 (0.061)	-0.006 (0.056)
Working hours per week	0.005 (0.004)	0.002 (0.004)	0.003 (0.004)
Fertility	0.040 (0.146)	-0.074 (0.153)	-0.073 (0.159)
Observations	6434		
Panel B: Girls			
Mother General Health - Excellent	-0.350 (0.331)	-0.434 (0.344)	-0.447 (0.333)
Mother General Health - Good	-0.301 (0.315)	-0.492 (0.325)	-0.357 (0.315)
Mother General Health - Fair	-0.380 (0.368)	-0.299 (0.392)	-0.083 (0.383)
Mother General Health - Poor	-0.100 (0.462)	-0.028 (0.507)	-0.181 (0.674)
Child longstanding illness	-0.045 (0.135)	0.223* (0.134)	0.089 (0.140)
Father longstanding illness	-0.001 (0.120)	-0.040 (0.119)	0.019 (0.130)
Income decile area	-0.019 (0.018)	-0.005 (0.019)	-0.011 (0.019)
Family and/or friends nearby - Yes friends	0.093 (0.241)	0.092 (0.250)	0.187 (0.245)
Family and/or friends nearby - Yes family	0.283 (0.309)	0.614 (0.309)	0.166 (0.326)
Family and/or friends nearby - Yes both	0.259 (0.229)	0.116 (0.237)	0.242 (0.232)
Non-parental child care	-0.004 (0.005)	-0.004 (0.006)	-0.005 (0.006)
Hours club	-0.023 (0.042)	0.003 (0.050)	-0.014 (0.056)
Working hours per week	-0.005 (0.004)	-0.002 (0.004)	-0.001 (0.004)
Fertility	0.090 (0.136)	0.020 (0.137)	-0.102 (0.143)
Observations	6470		

Notes: The table shows the full set of interactions of the benchmark model shown in Table 4. Data: UK Millennium Cohort Study, Main Sample.

Table A.11: Production model of socio-emotional skills (between age 6 and 11). Equivalence of self-productivity across socio-emotional traits.

	Differential effect of the input at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
relative to the 1st Quartile of socio-emotional skills			
Panel A: Boys			
Sensitive parenting style \leq median	-0.140 (0.091)	-0.311*** (0.093)	-0.282*** (0.098)
Sensitive parenting style $>$ median	0.253 (0.170)	0.508*** (0.168)	0.464*** (0.171)
Mother's mental health \leq median	-0.064 (0.068)	0.003 (0.062)	-0.098 (0.092)
Mother's mental health $>$ median	0.026 (0.197)	-0.145 (0.193)	-0.012 (0.217)
Time investment \leq median	0.054 (0.085)	-0.014 (0.084)	0.067 (0.088)
Time investment $>$ median	0.019 (0.153)	0.028 (0.155)	-0.090 (0.163)
Socio-emotional skills (age 6) (Internalising)		0.484*** (0.048)	
Socio-emotional skills (age 6) (Externalising)		0.447*** (0.045)	
σ^2 (variance of fixed effect)		0.468	
Observations		6434	
Panel B: Girls			
Sensitive parenting style \leq median	-0.013 (0.096)	-0.114 (0.108)	-0.240** (0.111)
Sensitive parenting style $>$ median	0.058 (0.162)	0.186 (0.169)	0.351** (0.170)
Mother's mental health \leq median	-0.214*** (0.058)	-0.202*** (0.070)	-0.189** (0.075)
Mother's mental health $>$ median	0.588*** (0.184)	0.594*** (0.196)	0.254 (0.199)
Time investment \leq median	-0.112 (0.091)	-0.092 (0.094)	-0.184* (0.097)
Time investment $>$ median	0.233 (0.163)	0.109 (0.161)	0.348** (0.167)
Socio-emotional skills (age 6) (Internalising)		0.293*** (0.050)	
Socio-emotional skills (age 6) (Externalising)		0.324*** (0.048)	
σ^2 (variance of fixed effect)		0.466	
Observations		6470	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study.

Table A.12: Production model of socio-emotional skills (between age 6 and 11). Controlling additionally for mother's and father's trait-specific socio-emotional skills

	Differential effect of the input at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
	relative to the 1st Quartile of socio-emotional skills		
Panel A:Boys			
Sensitive parenting style ≤ median	-0.162	-0.273**	-0.290***
	(0.100)	(0.101)	(0.107)
Sensitive parenting style > median	0.254	0.493**	0.480**
	(0.191)	(0.188)	(0.192)
Mother’s mental health ≤ median	-0.056	-0.019	-0.105
	(0.075)	(0.067)	(0.091)
Mother’s mental health > median	-0.005	-0.141	-0.022
	(0.216)	(0.210)	(0.234)
Time investment ≤ median	0.042	-0.016	0.061
	(0.094)	(0.092)	(0.097)
Time investment > median	0.013	0.014	-0.074
	(0.171)	(0.172)	(0.180)
Socio-emotional skills (age 6)		0.470***	
		(0.049)	
Mother’s Socio-emotional skills		0.057**	
		(0.024)	
Father’s Socio-emotional skills		0.024	
		(0.022)	
σ ² (variance of fixed effect)		0.458	
Observations		5430	
Panel B:Girls			
Sensitive parenting style ≤ median	0.015	-0.174	-0.244**
	(0.111)	(0.121)	(0.119)
Sensitive parenting style > median	0.038	0.234	0.335*
	(0.187)	(0.192)	(0.189)
Mother’s mental health ≤ median	-0.225***	-0.199**	-0.245**
	(0.075)	(0.083)	(0.095)
Mother’s mental health > median	0.598***	0.591***	0.398**
	(0.212)	(0.219)	(0.228)
Time investment ≤ median	-0.118	-0.066	-0.123
	(0.105)	(0.103)	(0.108)
Time investment > median	0.333*	0.125	0.348*
	(0.188)	(0.180)	(0.188)
Socio-emotional skills (age 6)		0.323***	
		(0.052)	
Mother’s Socio-emotional skills		0.013	
		(0.024)	
Father’s Socio-emotional skills		0.025	
		(0.022)	
σ ² (variance of fixed effect)		0.471	
Observations		5318	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, subsample with details on mother's and father's trait specific socio-emotional skills.

Table A.13: Descriptive statistics of the father's subsample, by gender

	Boys		Girls		Test diff
	mean	sd	mean	sd	p-value
Mother's internalising	0.003	1.010	-0.006	0.991	0.697
Mother's externalising	-0.013	1.018	0.011	0.983	0.316
Father's internalising	-0.013	0.993	0.015	1.008	0.305
Father's externalising	-0.038	0.994	0.037	1.006	0.005
Observations	2715		2659		5314

Notes: Last column indicates the p-value mean test of the difference in each variables between boys and girls. Data: UK Millennium Cohort Study, Father Sample.

Table A.14: Production model of socioemotional skills (between age 6 and 11). Including past inputs - Boys

	Differential effect of the input at the		
	(1) 2nd Quartile relative to the 1st Quartile of socio-emotional skills	(2) 3rd Quartile	(3) 4th Quartile
Boys			
Sensitive parenting style \leq median	-0.116 (0.096)	-0.293*** (0.097)	-0.303*** (0.105)
Sensitive parenting style $>$ median	0.224 (0.183)	0.462** (0.182)	0.473** (0.186)
Sensitive parenting style (age 3) \leq median	-0.006 (0.099)	-0.073 (0.098)	-0.007 (0.102)
Sensitive parenting style (age 3) $>$ median	-0.011 (0.167)	0.151 (0.170)	0.045 (0.173)
Mother's mental health \leq median	-0.053 (0.073)	0.022 (0.067)	-0.071 (0.099)
Mother's mental health $>$ median	0.112 (0.212)	-0.081 (0.213)	0.051 (0.241)
Mother's mental health (age 3) \leq median	0.007 (0.076)	0.005 (0.087)	-0.093 (0.101)
Mother's mental health (age 3) $>$ median	-0.294 (0.190)	-0.331 (0.205)	-0.141 (0.232)
Time investment \leq median	0.050 (0.089)	-0.030 (0.087)	0.072 (0.092)
Time investment $>$ median	0.013 (0.160)	0.026 (0.163)	-0.114 (0.172)
Time investment (age 3) \leq median	-0.203 (0.136)	-0.246* (0.136)	-0.256* (0.136)
Time investment (age 3) $>$ median	0.382* (0.206)	0.338* (0.202)	0.436** (0.206)
Socio-emotional skills (age 6)		0.479*** (0.046)	
σ^2 (variance of fixed effect)		0.475	
Observations		6148	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Standard errors are clustered at child level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, School Sample.

Table A.15: Production model of socioemotional skills (between age 6 and 11). Including past inputs - Girls

	Differential effect of the input at the		
	(1) 2nd Quartile relative to the 1st Quartile of socio-emotional skills	(2) 3rd Quartile	(3) 4th Quartile
Girls			
Sensitive parenting style \leq median	-0.011 (0.104)	-0.084 (0.116)	-0.241* (0.120)
Sensitive parenting style $>$ median	0.046 (0.171)	0.247 (0.178)	0.378* (0.182)
Sensitive parenting style (age 3) \leq median	0.011 (0.095)	-0.044 (0.097)	0.000 (0.000)
Sensitive parenting style (age 3) $>$ median	0.020 (0.157)	-0.154 (0.165)	-0.064 (0.166)
Mother's mental health \leq median	-0.178*** (0.068)	-0.204*** (0.076)	-0.167** (0.083)
Mother's mental health $>$ median	0.570*** (0.200)	0.548*** (0.210)	0.263 (0.217)
Mother's mental health (age 3) \leq median	-0.070 (0.082)	-0.073 (0.080)	-0.075 (0.097)
Mother's mental health (age 3) $>$ median	-0.118 (0.207)	0.127 (0.209)	-0.127 (0.220)
Time investment \leq median	-0.113 (0.094)	-0.080 (0.097)	-0.167* (0.100)
Time investment $>$ median	0.247 (0.168)	0.132 (0.164)	0.338** (0.172)
Time investment (age 3) \leq median	-0.185 (0.126)	-0.155 (0.126)	-0.207 (0.126)
Time investment (age 3) $>$ median	0.013 (0.178)	-0.066 (0.153)	0.000 (0.000)
Socio-emotional skills (age 6)		0.298*** (0.048)	
σ^2 (variance of fixed effect)		0.487	
Observations		6190	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Standard errors are clustered at child level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, School Sample.

Table A.16: Production model of socio-emotional skills between age 6 and 11. Combining mother's and teacher's reports to measure child socio-emotional skills (Sensitivity 1)

	Differential effect of the input at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
	relative to the 1st Quartile of socio-emotional skills		
Panel A: Boys			
Sensitive parenting style ≤ median	0.001	-0.221*	-0.202*
	(0.128)	(0.117)	(0.113)
Sensitive parenting style > median	0.038	0.342*	0.350*
	(0.213)	(0.204)	(0.191)
Mother’s mental health ≤ median	-0.150*	-0.018	-0.176**
	(0.082)	(0.079)	(0.087)
Mother’s mental health > median	0.271	0.062	0.134
	(0.234)	(0.231)	(0.241)
Time investment ≤ median	-0.073	-0.038	-0.002
	(0.104)	(0.099)	(0.102)
Time investment > median	0.145	-0.025	-0.010
	(0.195)	(0.186)	(0.192)
Socio-emotional skills (age 6)		0.413***	
		(0.056)	
σ ² (variance of fixed effect)		0.446	
Observations		4426	
Panel B: Girls			
Sensitive parenting style ≤ median	0.084	-0.058	-0.042
	(0.127)	(0.125)	(0.140)
Sensitive parenting style > median	-0.066	0.108	0.176
	(0.204)	(0.199)	(0.215)
Mother’s mental health ≤ median	-0.162**	-0.168**	-0.132
	(0.075)	(0.081)	(0.084)
Mother’s mental health > median	0.404*	0.384	-0.007
	(0.232)	(0.239)	(0.240)
Time investment ≤ median	0.059	0.051	-0.106
	(0.117)	(0.111)	(0.111)
Time investment > median	0.016	0.094	0.290
	(0.209)	(0.199)	(0.199)
Socio-emotional skills (age 6)		0.228***	
		0.059	
σ ² (variance of fixed effect)		0.462	
Observations		4413	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, subsample with details on teacher's reports of socio-emotional skills.

Table A.17: Production model of socio-emotional skills (between age 6 and 11), by gender. Allowing for differential self-productivity of socio-emotional skills across quartiles (Sensitivity 2).

	Differential effect of the input at the		
	(1)	(2)	(3)
	2nd Quartile	3rd Quartile	4th Quartile
relative to the 1st Quartile of socio-emotional skills			
Panel A: Boys			
Sensitive parenting style \leq median	-0.128	-0.299***	-0.272***
	(0.089)	(0.091)	(0.097)
Sensitive parenting style $>$ median	0.245	0.504***	0.466***
	(0.170)	(0.168)	(0.171)
Mother's mental health \leq median	-0.068	-0.001	-0.100
	(0.067)	(0.061)	(0.090)
Mother's mental health $>$ median	0.031	-0.142	-0.006
	(0.198)	(0.194)	(0.217)
Time investment \leq median	0.060	-0.011	0.068
	(0.084)	(0.083)	(0.087)
Time investment $>$ median	0.008	0.023	-0.093
	(0.151)	(0.155)	(0.162)
Socio-emotional skills (age 6 - first quartile)		0.473***	
	0.052	(0.053)	-0.105
	(0.124)	(0.170)	(0.100)
σ^2 (variance of fixed effect)		0.466	
Observations		6434	
Panel B: Girls			
Sensitive parenting style \leq median	-0.023	-0.121	-0.248**
	(0.096)	(0.108)	(0.110)
Sensitive parenting style $>$ median	0.065	0.184	0.354**
	(0.162)	(0.169)	(0.170)
Mother's mental health \leq median	-0.209***	-0.199***	-0.183**
	(0.057)	(0.069)	(0.074)
Mother's mental health $>$ median	0.595***	0.600***	0.256
	(0.185)	(0.197)	(0.200)
Time investment \leq median	-0.110	-0.092	-0.179*
	(0.091)	(0.093)	(0.096)
Time investment $>$ median	0.236	0.110	0.344**
	(0.162)	(0.160)	(0.166)
Socio-emotional skills (age 6) (age 6 - first quartile)		0.315***	
	-0.144	(0.053)	0.005
	(0.155)	(0.167)	(0.140)
σ^2 (variance of fixed effect)		0.466	
Observations		6470	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother's cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, Main Sample.

Table A.18: Production model of socio-emotional skills between age 6 and 11 by gender. Anchoring socio-emotional skills to ‘non-smoking’ at age 14 (Sensitivity 3).

	Differential effect of the input at the		
	(1) 2nd Quartile relative to the 1st Quartile of socio-emotional skills	(2) 3rd Quartile	(3) 4th Quartile
Panel A: Boys			
Sensitive parenting style \leq median	-0.104 (0.080)	-0.284*** (0.081)	-0.207** (0.088)
Sensitive parenting style $>$ median	0.147 (0.159)	0.401** (0.158)	0.310* (0.165)
Mother’s mental health \leq median	-0.059 (0.056)	0.075 (0.050)	-0.090 (0.075)
Mother’s mental health $>$ median	-0.075 (0.175)	-0.427* (0.164)	-0.095 (0.187)
Time investment \leq median	0.053 (0.071)	-0.038 (0.073)	0.038 (0.075)
Time investment $>$ median	-0.024 (0.139)	0.069 (0.140)	-0.049 (0.145)
Socio-emotional skills (age 6)		0.438*** (0.048)	
σ^2 (variance of fixed effect)		0.470	
Observations		5658	
Panel B: Girls			
Sensitive parenting style \leq median	0.012 (0.087)	-0.141 (0.097)	-0.183* (0.101)
Sensitive parenting style $>$ median	0.066 (0.144)	0.250 (0.155)	0.320** (0.157)
Mother’s mental health \leq median	-0.176*** (0.052)	-0.131** (0.064)	-0.146** (0.066)
Mother’s mental health $>$ median	0.508*** (0.167)	0.509*** (0.177)	0.189 (0.180)
Time investment \leq median	-0.053 (0.083)	-0.093 (0.086)	-0.135 (0.087)
Time investment $>$ median	0.139 (0.146)	0.107 (0.145)	0.261* (0.151)
Socio-emotional skills (age 6)		0.299*** (0.051)	
σ^2 (variance of fixed effect)		0.462	
Observations		5742	

Notes: The dependent variable socio-emotional skills at age 11 and the explanatory variable socio-emotional skills at age 6 are both expressed in standard deviations. All inputs are measured in standard deviations; with the exception of inputs that are dummy variables, income that is demeaned and expressed in £100 and mother’s cognitive skills which is the demeaned years of schooling. The full set of control variables are listed in Table 3. Estimated standard errors are robust to correlation in the error within child. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data: UK Millennium Cohort Study, subsample with details on child’s socio-emotional skills at age 14.