NBER WORKING PAPER SERIES

ESTIMATING THE TECHNOLOGY OF CHILDREN'S SKILL FORMATION

Francesco Agostinelli
Matthew Wiswall

Working Paper 22442
http://www.nber.org/papers/w22442

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2016, Revised March 2022

We thank the Editor, Jim Heckman, and four anonymous referees for their comments and feedback. We also thank Lance Lochner, Morteza Saharkhiz, Greg Veramendi, and participants at seminars and conferences at Arizona State, Barcelona GSE Summer Forum, Brown, CEHD (UChicago), Collegio Carlo Alberto, Duke, Interaction Workshop 2018 (Northwestern), Minnesota, Penn, Rice, University of Bristol, UCL, University of Houston, and Western Ontario for helpful comments. We are responsible for all errors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimating the Technology of Children's Skill Formation
Francesco Agostinelli and Matthew Wiswall
NBER Working Paper No. 22442
July 2016, Revised March 2022
JEL No. C38,J13

ABSTRACT

In this paper we study the process of children's skill formation. Using a dynamic latent factor structure, we show how measurement restrictions on observed measures aid the identification of skill technology features. We then use our identification results to develop and estimate the joint dynamic process of latent investment and skill development, allowing for static and dynamic complementarities in skill production between parental investments and children's skills. Using data for the United States, we estimate that parental investments are particularly productive in producing cognitive skills during early childhood (ages 5-6). Moreover, we find that the marginal productivity of investments in this period is substantially higher for children with lower existing skills, suggesting the optimal targeting of interventions to disadvantaged young children.

Francesco Agostinelli
University of Pennsylvania
133 South 36th Street
Philadelphia, PA 19104
fagostin@upenn.edu

Matthew Wiswall
Department of Economics
University of Wisconsin-Madison
1180 Observatory Drive
Madison, WI 53706
and NBER
matt.wiswall@gmail.com

An online appendix is available at http://www.nber.org/data-appendix/w22442

1 Introduction

The wide dispersion of measured human capital in children and its strong correlation with later life outcomes has prompted a renewed interest in understanding the determinants of skill formation among children (for a recent review, see Heckman and Mosso, 2014). However, the empirical challenges in estimating the skill formation process, principally the technology of child development, is hampered by the imperfect measures of children's skills we have available. While measurement issues exist in many areas of empirical research, they may be particularly salient in research about child development. There exists a number of different measures of children's skills, each measure can be arbitrarily located and scaled, and provide widely differing levels of informativeness about the underlying latent skills of the child. In the presence of these measurement issues, identification of the underlying latent process of skill development is particularly challenging, but nonetheless essential, because ignoring the measurement issues through ad hoc simplifying assumptions could bias the empirical conclusions.

This paper makes two contributions. First, using the estimation framework of Cunha and Heckman (2007, 2008) and Cunha et al. (2010) we develop alternative empirically grounded restrictions on the measurement of skills that allow identification of general skill technologies. Although our identification analysis considers only particular parametric classes of models and is less general than the non-parametric analysis in Cunha et al. (2010), the class of models we consider includes nearly all of the models researchers have estimated.[1] Second, we estimate a model of skill formation that allows for general interactions in skill production between a child's stock of existing skills and parental investments. This allows us to estimate the heterogeneity in the returns to parental investments at different stages of development and estimate sources of dynamic complementarities between early and later investments.

We analyze the concept of "age-invariant" measures, measures that allow the comparison of skill development as children age and imply restrictions on how measures relate to each other over the development period. These assumptions are certainly not appropriate for all measures, but at least some skill measures are designed by psychometricians and child development researchers for these kinds of purposes. We show that if these types of measures are available, then more general skill technolo-

[1] Our particular point of departure is the log-linear measurement models estimated in Cunha et al. (2010) and other recent papers (e.g., Attanasio et al., 2020). There are only a few exceptions to this in the current literature that we are aware of. See Del Boca et al. (2014a) and Williams (2019) for examples of studies that allow for discrete measures, and Bond and Lang (2013, 2018) for an example of non-linear transformations of test scores. Cunha et al. (2021) provide a recent summary of this research.

gies can be identified. But if these measures are unavailable, then one can consider alternative restrictions, in particular some restriction on the skill formation technology. There is therefore a tradeoff between restrictions on the measurement model and the skill technology, where restrictions on one reduce the need for restrictions on the other.

We also provide some guidance for empirical researchers. We show in particular that age-standardizing measures (Z-scores), a popular technique for transforming and describing skill distributions does not guarantee comparability as required for identifying the underlying skill technology. As we show, age-invariant measures require assumptions about the relationship between unobserved skills and observed measures, which can neither be directly tested nor guaranteed through generic data transformations.

We then estimate a flexible parametric version of our model using data from the US National Longitudinal Survey of Youth (NLSY). We examine the development of cognitive skills in children from age 5 to age 14, and estimate a model of skill development allowing for complementarties between parental investment and children's skills; endogenous parental investment responding to the stock of children's skills, maternal skills, and family income; Hicks neutral dynamics in TFP and free returns to scale; and unobserved shocks to the investment process and skill production. Following Cunha et al. (2010), our empirical framework treats not only the child's cognitive skills as measured with error, but investment and maternal skills as well.

Following our identification analysis, we develop a multiple step instrumental variable estimator. Our estimator is not only relatively simple and tractable, but also robust to parametric distributional assumptions on the marginal distribution of latent variables and measurement errors, as is commonly imposed in the prior empirical literature. We jointly estimate the technology of skill formation, the process of parental investments in children, and the adult distribution of completed schooling and earnings. We also allow the parameters of the production technology and investment process to freely vary as the child ages. The measures in the NLSY dataset of cognitive achievement for children (PIAT scores), which were designed to account for developmental changes in children's skills, are assumed to be age-invariant over the age range we consider (ages 5-14). Our estimates of high TFP and increasing returns to scale at early ages indicate that investments are particularly productive early in the development period. In contrast to existing estimates of positive static complementarity in skills, we find that the marginal productivity of early investments is substantially higher for children with *lower* existing skills, suggesting the optimal targeting of interventions to disadvantaged children.

Our estimates of the dynamic process of investment and skill development allow us to estimate the heterogeneous treatment effects of some simple policy interventions. We show that income transfer to families at age 5-6 would substantially increase children's skills and completed schooling, with the effects larger for low income families. When we compare these estimates to those using models that ignore measurement error, we estimate policy effects which are substantially smaller, indicating that the methods we adopt are important quantitatively to answering key policy questions.

The next three sections present our empirical model, identification analysis using a simplified model, and estimation procedure including a discussion of data sources and measures. The remaining sections present our empirical results, comparisons to existing estimates, and policy counterfactual results.

2 Stylized Model

In this Section, we describe our model of skill development. Here, we omit several model features to simplify the discussion and subsequent identification analysis in the next section. We introduce these elements in later sections when we take the model to the data.

2.1 Skill Production Technology

Child development takes place over a discrete and finite period, $t = 0, 1, \ldots, T$, where $t = 0$ is the initial period (say birth) and $t = T$ is the final period of childhood (say age 18). There is a population of children and each child in the population is indexed i. For each period, each child is characterized by a skill stock $\theta_{i,t}$ and a flow investment $I_{i,t}$. In what follows, we consider only a single scalar skill and scalar investment, but it is straight-forward to generalize this to multiple skills and multiple investments. For each child, the current stock of skill and current flow of investment produce next period's stock of skill according to the skill formation production technology:

$$\theta_{i,t+1} = h_t(\theta_{i,t}, I_{i,t}, \eta_{i,\theta,t}) \text{ for } t = 0, 1, \ldots, T-1 \tag{1}$$

where $\eta_{i,\theta,t}$ is a production shock. Equation (1) can be viewed as a dynamic state space model with $\theta_{i,t}$ the state variable for each child i. The production technology $h_t(\cdot)$ is indexed with t to emphasize that the technology can vary as children age. According to this technology, the sequence of investments and shocks and the initial stock of child skills $\theta_{i,0}$ produce the sequence of skill stocks for each child i: $\theta_{i,0}, \theta_{i,1}, \ldots, \theta_{i,T}$.

In the remainder of the paper, we work with technologies of the form:

$$\ln \theta_{t+1} = \ln A_t + \psi_t \ln f_t(\theta_t, I_t) + \eta_{\theta,t} \qquad \text{for } t = 0, 1, \ldots, T-1\,, \tag{2}$$

where, we have dropped the i subscript to simplify our notation. $f_t(\theta_t, I_t)$ represents a production function whose location and scale (in logs) are fully represented by the total factor productivity (TFP) term A_t and the return to scale parameter ψ_t, respectively.[2] $f_t(\theta_t, I_t)$ can be specified as a constant return to scale Cobb-Douglas function, a more general CES function, or various "trans-log" functions, among others.[3]

2.2 Policy-Relevant Effects of Interest

There are several features of the technology which have particular relevance to understanding the process of child development and in evaluating policy interventions to improve children's skills. First, a key question is the productivity of investments at various child ages. At what ages are investments in children particularly productive in producing future skills ("critical periods") and, conversely, at what ages is it difficult to re-mediate deficits in skill? Second, how does heterogeneity in children's skills, at any given period, affect the productivity of new investments in children? Complementarity in the production technology between current skill stocks and investments implies heterogeneity in the productivity of investments across children. Third, how do investments in children persist over time and affect adult outcomes? Do early investments have a high return because they increase the productivity of later investments (dynamic complementarities) or do early investments "fade-out" over time? These features of the technology of skill development then directly inform the optimal *timing* of policy interventions – the optimal investment portfolio across early and late childhood – and the optimal *targeting* of policy – to which children should scarce resources be allocated to, with the goal of using childhood interventions to affect eventual adult outcomes.

[2]In the Appendix, in a more formal semi-parametric analysis, we define the function f_t as one with known location and scale, i.e. the location and scale parameters for the h_t super-function (1) are represented by the $\ln A_t$ and ψ_t parameters, and the sub-function f_t therefore has no unknown parameters for scale and location.

[3]See Cunha et al. (2010) for a non-parametric analysis of more general technologies.

2.3 Complementarities and Heterogeneity in Skill Production

We conclude this Section by emphasizing the importance of functional forms to the policy questions of interest described above, and in particular the heterogeneity in the productivity of parental investments. Much of the previous empirical work adopted parametric specifications for the f_t function assuming CES forms (general or Cobb-Douglas special cases), see for example Cunha and Heckman (2008); Cunha et al. (2010). In these cases, the marginal return to parental investments is assumed to be (weakly) positive with respect to the current stock of skills:

$$\frac{\partial^2 \theta_{t+1}}{\partial I_t \partial \theta_t} \geq 0 \; \forall t.$$

Although this assumption of a positive contemporaneous complementarity is generally non-controversial when the inputs are labor and capital, in this case it implies a specific pattern of heterogeneous marginal products in the current period: the marginal product of parental investments is larger for higher skilled children. Although, as shown by Cunha and Heckman (2008) and Cunha et al. (2010), even with this type of function, it still may be optimal in terms of later adult outcomes to target investments to skill-disadvantaged children at early ages because of a dynamic complementarity.

In our empirical model, we consider a simple alternative model utilizing trans-log forms that allow for both positive and negative contemporaneous/static complementarities:

$$\ln \theta_{t+1} = \ln A_t + \gamma_{1,t} \ln \theta_t + \gamma_{2,t} \ln I_t + \gamma_{3,t} \ln I_t \cdot \ln \theta_t + \eta_{\theta,t}, \tag{3}$$

where $\gamma_{3,t}$ is a free parameter that characterizes the heterogeneity in the returns to parental investments. A positive value of $\gamma_{3,t}$ indicates that parental investments are more productive for already highly skilled children, and a negative value implies a higher productivity for skill disadvantaged children. Although it does not nest the CES form, this specific parametric model is flexible in the sign of the static complementarity.

3 Measurement and Identification

In this Section we discuss our main identification result. In particular, we show that even with common assumptions about measurement errors, production function parameters and key policy relevant effects are under-identified without further restrictions. A key issue is that *changes* over the development period in the location and scale of the production function cannot be identified separately from *changes* in the location and scale of the measures. We discuss a range of restrictions that are sufficient for identification, and focus in particular on empirically grounded restrictions on the measurement process that allow identification of more general technologies.

3.1 Measurement Model

The focus of this paper and much of the recent literature in economics is estimating the technology determining child skill development (1), while accommodating the reality that researchers have at hand various often arbitrarily constructed and imperfect measures of children's skills. Following the influential prior work in this area (Cunha and Heckman, 2007; Cunha et al., 2010; Cunha and Heckman, 2008), our framework recognizes that children's skills are not directly measured by a single measure, but there exists multiple measures, which can have some relationship to the unobserved latent skill stock θ_t.

We follow the vast majority of the empirical literature and assume a (log) linear system of measures, which is certainly not without a loss of generality as we discuss below.[4] For each period t, we have M_t measures for latent skills $\ln \theta_t$: $m = 1, 2, \ldots, M_t$. $Z_{t,m}$ represents the specific observed measure, and it is modeled as:

$$Z_{t,m} = \mu_{t,m} + \lambda_{t,m} \ln \theta_t + \epsilon_{t,m} \qquad \text{for } t = 0, 1, \ldots, T \text{ and } m = 1, \ldots, M_t. \tag{4}$$

The measurement parameters $\mu_{t,m}$ and $\lambda_{t,m}$ represent the location and scale of the measures, respectively, and $\epsilon_{t,m}$ is the measurement error, with $E(\epsilon_{t,m}) = 0$ without loss of generality. We assume measurement errors are independent of latent skills and investment and independent of each other at each period t. To simplify the identification analysis in this Section, we assume that investments are observed. Following this identification analysis, we present a more general model relaxing this assumption, and discuss specific assumptions sufficient for identification of the model with unobserved investment, which are straight-forward extensions of the analysis presented here.

[4] For example, the empirical analysis of Cunha and Heckman (2007); Cunha et al. (2010); Cunha and Heckman (2008) and Attanasio et al. (2020).

As has been long recognized, measures of child skills (e.g. test scores and the like) are fundamentally ordinal: they indicate a rank ordering of children on the particular measure, but do not necessarily provide an interval or cardinal representation of latent skills (see Cunha et al. (2021) for a recent review). However, the log-linear measurement model we use (4), following much of the current literature, does impose a kind of cardinality on the measures in that values of the skill measures (in expectation) map into specific values of latent log skills, up to the unknown location and scale parameters $\mu_{t,m}, \lambda_{t,m}$. A number of papers have discussed the problems with such assumptions and various solutions.[5]

While acknowledging this important issue, our focus here is on a separate issue: how the measures relate to each other across periods/ages (t). We show that this is critical to identifying the dynamics in the production technology, and in particular how the productivity of investments change as children age. And, moreover, this issue would still be a relevant issue even with more general measurement models (ones that allow for non-linear or discrete relationships).

3.2 Under-Identification Problem

We begin by noting that with some normalization in the initial period ($t = 0$), and at least 3 measures of skills in this period, we identify the distribution of initial latent skills.[6] But without further restrictions, we cannot separately identify the location and scale of the measures in periods after the initial one from the scale and location of the production technology. Some intuition comes from analyzing the change in the mean of the observed measures between $t + 1$ and t:

$$\begin{aligned} E(Z_{t+1,m}) - E(Z_{t,m}) &= (\mu_{t+1,m} - \mu_{t,m}) + \lambda_{t+1,m} E(\ln \theta_{t+1}) - \lambda_{t,m} E(\ln \theta_t) \\ &= \underbrace{(\mu_{t+1,m} - \mu_{t,m}) + (\lambda_{t+1,m} - \lambda_{t,m})}_{\text{measurement}} E(\ln \theta_t) + \lambda_{t+1,m} \underbrace{(E(\ln \theta_{t+1}) - E(\ln \theta_t))}_{\text{latent skills}} \end{aligned}$$

where an increase in the average observed measures can be attributed to either a change in the measurement parameters such that the measures have become "easier" (the first two terms) or to an increase in average latent skills (the last term).

[5] See for example Bond and Lang (2013, 2018) for issues related to measuring black-white skill gaps, and Ballou (2009) for issues in the education value-added context. See Del Boca et al. (2014a) and Williams (2019) for examples of studies that allow for discrete measures. Cunha et al. (2021) discuss more examples.

[6] This is a well known result and details for our full model are provided in the next Section. Repeated "re-normalizations" for each period imply restrictions on the parametric model, and are not in general normalizations, as discussed in Agostinelli and Wiswall (2016).

An analogous under-identification problem exists when attempting to infer changes in the productivity of investments as children age from changes in the marginal product of mean measures with respect to investment:

$$\frac{\partial E(Z_{t+1,m}|I_t)}{\partial I_t} - \frac{\partial E(Z_{t,m}|I_{t-1})}{\partial I_{t-1}} = \lambda_{t+1,m}\frac{\partial E(\ln\theta_{t+1}|I_t)}{\partial I_t} - \lambda_{t,m}\frac{\partial E(\ln\theta_t|I_{t-1})}{\partial I_{t-1}}$$
$$= \underbrace{(\lambda_{t+1,m} - \lambda_{t,m})}_{\text{measurement}}\frac{\partial E(\ln\theta_t|I_{t-1})}{\partial I_{t-1}} + \lambda_{t+1,m}\underbrace{\left(\frac{\partial E(\ln\theta_{t+1}|I_t)}{\partial I_t} - \frac{\partial E(\ln\theta_t|I_{t-1})}{\partial I_{t-1}}\right)}_{\text{latent productivity}}$$

where recall that our timing convention specifies θ_{t+1} a function of I_t investment, and we further assume to simplify the analysis that investment is independent of the measurement errors and production shock. Here an increase in the partial derivative of average skill measures with respect to investment can either be attributed to a change in the measurement scale parameters as the measures become more "sensitive" to latent skills (first term) or to a change in the primitive technology as investments become truly more productive in producing latent skills (last term). The lack of identification implies that without further restrictions we cannot simply use changes in the sensitivity of measures to investment to infer "sensitive" or "critical" periods, as in the Cunha and Heckman (2008) analysis, periods in which children's skills are particularly malleable and interventions are likely to be particularly effective.

This under-identification issue has a close analog in the extensive literature estimating teacher "value-added" using student test scores (Chetty et al., 2014). Rare in this literature is data allowing teacher quality to be anchored to adult outcomes (for an exception see Chetty et al., 2011), and almost the entirety of the empirical results rest on particular measures of student skills, typically grade-specific standardized test scores (for some discussion of general issues see Ballou, 2009). In the Appendix, we provide a value-added framework connecting our child development model to the standard estimating equations in the education value-added literature, where we replace our generic childhood investments with teacher assignment.[7] We conclude that the main statistic of interest in this literature–the variance of the estimated teacher effects, which indicates how "productive" teachers are in affecting student learning–is not a scale-free parameter. Much as in the previous case, the variance in the teacher effects estimated in one grade versus another can vary widely, not only

[7]Note that the same issue applies to estimating the productivity of other education inputs, including school or school district level value-added. For recent work estimating latent factor models including both parental and school inputs, see Agostinelli et al. (2019).

because of changes in the underlying productivity of teachers across grades, but because of changes in the measures. Although the common scaling parameter does not affect the *rankings* of teachers within grade, if the measurement scale changes across grades, we cannot use standard value-added estimates to infer the effect of policies that would for example reallocate teachers across grades, the analogous policy to re-distributing generic childhood investments from late to early periods as analyzed here and elsewhere, by Cunha et al. (2010) for example.

In the remaining analysis we focus on the identification of the semi-parametric form given in (2):

$$\ln \theta_{t+1} = \ln A_t + \psi_t \ln f_t(\theta_t, I_t) + \eta_{\theta,t} \qquad \text{for } t = 0, 1, \ldots, T-1$$

The previous empirical literature has primarily focused on standard parametric specifications for the technology of skill formation, such as Cobb-Douglas and CES specifications for the f_t sub-function, with constant return to scale $\psi_t = 1$ and unitary TFP $A_t = 1$ for all t.[8]

The technology (2) has a free location and scale represented by the $\ln A_t$ parameter (TFP in levels) and the ψ_t parameter (returns to scale in levels), respectively. Allowing for the log-linear measurement model (4), these technology parameters are not identified due to the missing "link" between skill measures across periods. Combining equations (4) and (2) we have

$$Z_{t+1,m} = \underbrace{(\mu_{t+1,m} + \lambda_{t+1,m} \ln A_t)}_{= \beta_{0,t,m}} + \underbrace{(\lambda_{t+1,m}\psi_t)}_{= \beta_{1,t,m}} \ln f_t(\theta_t, I_t) + u_{t,m}, \qquad (5)$$

where the new error term is $u_{t,m} = \lambda_{t+1,m}\eta_{\theta,t} + \epsilon_{t+1,m}$. Equation (5) indicates that even if the "reduced-form" parameters $(\beta_{0,t,m}, \beta_{1,t,m})$ are identified, the technology location and scale parameters (A_t, ψ_t) are not separately identified from the next period measurement location and scale parameters $(\mu_{t+1,m}, \lambda_{t+1,m})$. Next, we provide a framework that exploits the design of specific skill measures to separately identify these features.

3.3 Age-Invariance

An extensive literature, principally in psychometrics and education, is concerned with designing skill measures that can be vertically scaled across children of different

[8] See for example the estimated models in Cunha and Heckman (2008) and Cunha et al. (2010). Cunha et al. (2010) has non-parametric identification results for more general models.

ages so that the development of children can be consistently tracked.[9] These types of measures primarily consist of tests designed to be applicable for children of various ages, and include a range of test items which show meaningful variation for both younger and older children. An example is a vocabulary test in which children are asked to define words of increasing difficulty. To the extent that one could administer this same test to children in a range of ages, the count of words defined correctly could be considered an age-invariant measure of latent skill in this domain.

In our framework, we formalize this idea by defining *age-invariant* measures:

Definition 1 *A pair of measures* $Z_{t,m}$ *and* $Z_{t+1,m}$ *is* age-invariant *if* $E(Z_{t,m}|\theta_t = p) = E(Z_{t+1,m}|\theta_{t+1} = p)$ *for all* $p \in \mathbb{R}_{++}$.

Age-invariant measures imply that two children with the same level of latent skill would on average perform equally well, independently of their age. Age-invariance can be thought of as a kind of factor model restriction, akin to those commonly imposed limiting the number of latent skills. In this case, the restriction is that the only relevant variable for the measure is the level of latent skill possessed by the child, not the child's age directly. Definition 1 together with the assumptions on our measurement model implies that the measurement parameters for a specific age-invariant measure m are constant over the two age periods ($\mu_{t,m} = \mu_{t+1,m}$ and $\lambda_{t,m} = \lambda_{t+1,m}$).[10]

Whether a given pair of measures is age-invariant depends on the measures available, and must be evaluated on a case-by-case basis. Using pairs of unrelated measures, such as counts of body parts a toddler can identify to measure skills at age 1 and SAT scores to measure skills at age 18, would not constitute a pair of age-invariant measures as there is no reason to believe they would even have the same

[9]See the extensive literature review in Kolen and Brennan (2014). They review the methods to develop vertically scaled measures and various challenges. In their taxonomy, the measures we use in our empirical exercise were developed using a "common item design" (in which students at different ages receive at least some overlapping test items, as discussed in more detail below). They also review the extensive literature evaluating these methods and various ad hoc ways researchers and practitioners have attempted to construct vertical scales from measures not necessarily designed for this purpose.

[10]Age-invariance, together with the linear measurement model and the mean-independence assumption of the measurement noise ($E[\epsilon_{t,m}|\ln\theta_t] = E[\epsilon_{t,m}] = 0$), implies the following restrictions on measurement parameters for our log-linear measurement model: $\mu_{t+1} + \lambda_{t+1}\ln p = \mu_t + \lambda_t \ln p$ for all p. Re-arranging, we have $(\mu_{t+1} - \mu_t) = \ln p\,(\lambda_t - \lambda_{t+1})$ for all p. This is the case if and only if $\mu_t = \mu_{t+1}$ and $\lambda_t = \lambda_{t+1}$. Note that given our restricted measurement model, it would be sufficient to assume this condition directly. However, we prefer to work with the more general definition of age-invariance to build intuition.

location and scale. Other measures may be age-invariant, such as certain test score measures developed specifically to track development as children age. Examples of these types of measures for the cognitive skill domain include the Peabody Individual Achievement Test (PIAT) and the Woodcock-Johnson tests, both of which are specifically designed to track child development over a wide range of child ages. These measures have have been used in numerous studies, both inside and outside economics. In our empirical application, we use the PIAT measures and discuss in more detail why we believe these measures are age-invariant.[11]

A few remarks:

- Age-invariance is not required across all periods. A sufficient condition for identification of the technologies considered here is that there are at least some pairs of measures ages (e.g., from 5 to 6, from 7 to 8, etc.) that are age-invariant for these age-pairs. Moreover, age-invariance is not required for all measures, and additional non-age-invariant measures can be incorporated along with the age-invariant ones.

- Age-invariance does not impose any particular correlation between measures across ages. Age-invariance is not an assumption that measures are equally informative at all ages. There can be various degrees of measurement noise in individual measures such that there is low inter-period/age correlations in age-invariant measures. A corollary is then that a small inter-age correlation in measures does not necessarily suggest a lack of age-invariance, although it may suggest that the skill domains measured are changing across ages (see e.g. Lewis and McGurk, 1972).

- More generally, age-invariance does not impose any particular structure on the marginal distribution of latent skills nor on the measurement errors. Age-invariance for example does not rule out ceiling or floor effects whereby the relationship between latent skills and measures is non-linear. Following much of the existing literature we impose log-linearity in the measurement system and assume latent skills and measurements errors are independent, but it is

[11]Several recent papers explicitly invoke age-invariance assumptions regarding their particular measures. Attanasio et al. (2020) argue that the raw count of number of tasks completed from the Bayley cognitive scale is age-invariant over the age range they consider. Using the same data, Attanasio et al. (2019a) argue that a transformed age-equivalent version of the Bayley is age-invariant. And, Attanasio et al. (2019b) use the Peabody Picture Vocabulary Test (PPVT) as an age-invariant measure. These studies also assume age-invariance for certain non-cognitive and health measures. As in our paper, these assumptions are justified on a case-by-case basis.

possible to use the primitive age-invariance assumption in combination with other measurement models, including those allowing for discrete or non-linear measures (e.g., Williams, 2019).[12]

- Age-invariant measures need not imply overlap in the skill distribution across ages. It is possible that in some population there are no two children of different ages with the same level of latent skills. Age-invariance imposes general restrictions on the measures, in particular the common scale and location in our log-linear case.

- Our concept of age-invariance should not be confused with "age-standardized" or "age-equivalent" measures. The latter are two transformations of the raw data: age-standardized measures are constructed to be mean 0 and standard deviation 1 at each age, and age-equivalent measures are constructed to express raw scores relative to the typical development pattern using mean or median scores by age. Although age-equivalence transformations are actually predicated on the assumption of age-invariance, they do not guarantee that the resulting transformed measures are age-invariant. Age-invariant measures cannot be automatically constructed using *ex post* data transformations because the age-invariance property concerns the relationship between data and unobserved latent skills, which must be argued for on a case-by-case basis. However, it is possible that certain transformations may result in measures which are age-invariant whereas the raw measures are not, and vice versa.

  Note also that our concept of age-invariance is unrelated to the concept of "anchoring" (Cunha et al., 2010). Anchoring in this influential work is a transformation of the latent variables in terms of adult outcomes (e.g. earnings). Age-invariance, and the related measurement parameter assumptions, concern the relationship between skill measures and their associated latent variables during childhood.

- In practice, age-invariant measures need not only encompass measures with testing instruments that are given exactly to children of different ages (i.e. children given the exact same test questions). Many tests intended to track development across ages are often administered such that the questions are endogenously determined by the previous answers of the child. Therefore, while not all children are in fact answering the exact same test questions, their scores

[12]We do note that substantial floor or ceiling lumping evident in the raw data may cast doubt on the age-invariance assumption. As discussed below, our particular assumed age-invariant measures do not exhibit substantial floor or ceiling lumping.

are determined in an age comparable way. The typical test includes a number of test items ranging from low to high difficulty questions. Testing begins by first establishing a baseline test item for each child. While the baseline is initially based on the child's age, the baseline is adjusted downward (to less difficult questions) as the child is unable to answer questions correctly. Once the baseline is established, the test then progressively asks more difficult questions. Testing stops when the child makes a certain number of mistakes. The score is then determined as the number of correct answers before testing stops. Included in this number of correct answers are the lower difficulty test items prior to the baseline item because it is assumed the child would have answered these items correctly (given she was able to answer more the difficult items).

- Finally, we note that while we work with a sufficient condition for measures such that we exclude age from the measurement system, we could weaken this exclusion condition to allow some restricted parametric forms of age-varying in the measures. For example, if the researcher believes in some constant linear age-varying profile in the measures, this could be incorporated into the measurement system and identification of the primitive technology likely would still obtain. We leave these hybrid cases to future work.

3.4 Cobb-Douglas Example

In this final identification section we provide one simple parametric example illustrating how age-invariant measures can identify a simple production technology. We continue to work with a simplified model in which we assume investment is perfectly observed and exogenous. In the next section we analyze the full model we bring to data, and the Appendix provides a more general semi-parametric identification analysis.

Consider a Cobb-Douglas production function, which for skills developed in period 1 is specified as:

$$\ln \theta_1 = \ln A_0 + \psi_0(\gamma_0 \ln \theta_0 + (1 - \gamma_0) \ln I_0) + \eta_{\theta,0} \tag{6}$$

with $\gamma_0 \in (0, 1)$. The parameters $\ln A_0$ and ψ_0 represent the location and scale of the production function.

Assuming that the initial period ($t = 0$) measurement parameters $\{\mu_{0,m}, \lambda_{0,m}\}_m$ are already identified, we define the "error-contaminated" measures for the initial period and the next period measures as:[13]

$$\widetilde{Z}_{0,m} \equiv \frac{Z_{0,m} - \mu_{0,m}}{\lambda_{0,m}} = \ln \theta_0 + \widetilde{\epsilon}_{0,m} \ ,$$

where $\widetilde{\epsilon}_{0,m} = \frac{\epsilon_{0,m}}{\lambda_{0,m}}$, while the next period measure is a function of latent skills as follows:

$$Z_{1,m} = \mu_{1,m} + \lambda_{1,m} \ln \theta_1 + \epsilon_{1,m}.$$

As in all of our analysis above, the measurement parameters $\mu_{1,m}$ and $\lambda_{1,m}$ are treated as free parameters. Substituting the production technology into the period 1 measurement equation, we have

$$\begin{aligned} Z_{1,m} &= \mu_{1,m} + \lambda_{1,m} \ln A_0 + \lambda_{1,m}\psi_0(\gamma_0 \ln \theta_0 + (1-\gamma_0)\ln I_0) + \lambda_{1,m}\eta_{\theta,0} + \epsilon_{1,m} \\ &= (\mu_{1,m} + \lambda_{1,m} \ln A_0) + \lambda_{1,m}\psi_0(\gamma_0(\widetilde{Z}_{0,m} - \widetilde{\epsilon}_{0,m}) + (1-\gamma_0)\ln I_0) + \lambda_{1,m}\eta_{\theta,0} + \epsilon_{1,m} \\ &= \beta_{0,0} + \beta_{0,1}\widetilde{Z}_{0,m} + \beta_{0,2} \ln I_0 + \pi_{0,m} \end{aligned} \tag{7}$$

where equation (7) is a reduced-form equation of the original technology, with

$$\pi_{0,m} = \lambda_{1,m}\eta_{\theta,0} + \epsilon_{1,m} - \lambda_{1,m}\psi_0\gamma_0\widetilde{\epsilon}_{0,m}.$$

The reduced-form parameters $\beta_{0,0}, \beta_{0,1}, \beta_{0,2}$ are combinations of unknown measurement parameters $\mu_{1,m}, \lambda_{1,m}$ and unknown production function parameters $\ln A_0, \psi_0, \gamma_0$:

$$\beta_{0,0} = \mu_{1,m} + \lambda_{1,m} \ln A_0,$$

$$\beta_{0,1} = \lambda_{1,m}\psi_0\gamma_0,$$

$$\beta_{0,2} = \lambda_{1,m}\psi_0(1-\gamma_0)$$

Assuming that the reduced-form parameters (βs) are identified, we are still faced with an under-identification problem as there are 5 unknown primitive parameters $\ln A_0, \psi_0, \gamma_0, \mu_{0,m}$, $\lambda_{1,m}$, and there are only 3 identified reduced form parameters in

[13] In the estimation section we describe how to identify the initial measurement parameters and initial latent skill distribution. Our identification of the initial conditions follows standard arguments used in the current literature (e.g., Cunha et al., 2010).

(7). One could impose a technology restriction (e.g. $\ln A_0 = 0$ and $\psi_0 = 1$), but our goal is to use the available data to identify a general technology.[14]

We assume that the measures $Z_{0,m}, Z_{1,m}$ constitute a pair of age-invariant measures (Definition 1). This implies the measurement restriction $\mu_{0,m} = \mu_{1,m} = \mu_m$ and $\lambda_{0,m} = \lambda_{1,m} = \lambda_m$. Given the identification of the initial period measurement parameters $\mu_{0,m}, \lambda_{0,m}$, we can then identify the remaining technology parameters:

$$\psi_0 = \frac{\beta_{0,1}+\beta_{0,2}}{\lambda_m},$$

$$\gamma_0 = \frac{\beta_{0,1}}{\beta_{0,1}+\beta_{0,2}},$$

$$\ln A_0 = \frac{\beta_{0,0}-\mu_m}{\lambda_m}.$$

We can continue in this way, and with the presence of age-invariant measures for the remaining periods, to identify the full sequence of primitive technology parameters.[15]

4 Estimation

In this Section, we move beyond the stylized model analyzed in preceding sections and lay out the specific functional forms of the more general model we estimate. Our specification of the multi-dimensional initial conditions, parental investment functions, and adult outcome equations largely follow the empirical model of Cunha et al. (2010).[16] We conclude this Section by discussing the estimation algorithm:

[14]These kinds of production function restrictions can also be understood as restrictions in a traditional "reduced form" error-in-variables model (Chamberlain, 1977). In this literature, identification is often achieved by a proportionality restriction (e.g. linear regression parameters are assumed proportional to each other), i.e. restrictions imposed on the β parameters directly. In our case, the restrictions we consider come from restrictions on the primitive production function or measurement system, which is intuitively appealing because we can understand the consequences of these restrictions more clearly.

[15]Some more recent studies have provided complementary identification results in similar settings. For example, Del Bono et al. (2020) show that if researchers are interested in the average marginal effects of improving childhood inputs on later in life adult outcomes, this only requires the identification of reduced form treatment effects, and not separate identification of the technology and measurement parameters. In addition, Freyberger (2020) provides several additional identification results, including new results on the identification of CES production functions.

[16]Two notable differences: First, we do not model the child's non-cognitive skill production. Second, we do not assume a particular marginal distribution for the latent variables or measurement errors (assumed Normal or mixture of two Normal distributions in Cunha et al., 2010), with the exception of the initial conditions distribution.

a sequential instrumental variable estimator with instruments derived from primitive assumptions about the correlation of measurement errors, latent variables, and unobserved shocks.

4.1 Functional Forms

4.1.1 Multidimensional Initial Conditions

We assume periods t last for 2 years, and the initial period $t = 0$ corresponds to age 5-6 and the last period $t = 4$ corresponds to ages 13-14. The full set of initial conditions consist of the child's initial (at age 5-6) stock of skills θ_0, the mother's cognitive and non-cognitive skills (θ_{MC} and θ_{MN}), which are assumed to be time invariant over the child development period, and the level of family income at the initial period (Y_0). Define the vector of initial conditions as

$$\Omega = (\ln \theta_0, \ln \theta_{MC}, \ln \theta_{MN}, \ln Y_0)$$

Although our estimation algorithm for the skill technology does not require any parametric specification for the initial conditions, for counterfactual exercises we need to model and estimate the initial relationship between latent skills and family income. For this reason, we assume a parametric distribution for the initial conditions:

$$\Omega \sim N(\mu_\Omega, \Sigma_\Omega)$$

where $\mu_\Omega = [0, 0, 0, \mu_{0,\ln Y}]$. $\mu_{0,\ln Y}$ is the mean of log household income when children are 5-6 years old. The means of the remaining variables are set to zero as a normalization. Σ_Ω is the variance-covariance matrix for the initial conditions. With the exception of income, the remaining variables are treated as latent factors with measures described below.

4.1.2 Parental Investments

We specify a parametric policy function for parental investment. The parametric policy function depends on the current stock of the child's skills, mother's skills, and family income:

$$\ln I_t = \alpha_{1,t} \ln \theta_t + \alpha_{2,t} \ln \theta_{MC} + \alpha_{3,t} \ln \theta_{MN} + \alpha_{4,t} \ln Y_t + \eta_{I,t} \tag{8}$$

Y_t is household income in period t, and $\eta_{I,t}$ is the investment shock, where $\eta_{I,t}$ i.i.d. $\sim N(0, \sigma^2_{I,t})$ for all t, and is assumed independent of latent skills and income. Moreover, we assume that the investment policy function has constant returns to scale, where

$\sum_j \alpha_{j,t} = 1$ for all t, following Cunha et al. (2010); Attanasio et al. (2015a,b). Because the set of available measures of parental investments change with the child's age, we cannot assume that the investment measures are age-invariant, which implies that the scale of the investment equation cannot be separately identified from the age-varying measurement factor loadings. Finally, our concept of investment represents both quantity and quality aspects, where we use measures of investments which capture quantity aspects (e.g. time parents spent reading to children) and quality aspects (e.g. whether children are "praised" by their parents).

This specification of investment is an approximation of the parental behavior, and is not derived from an explicit economic model of the household behavior. The advantages of this approach are twofold. First, it provides a simple and tractable model of the investment process, which avoids the computational burden of solving and estimating a formal model of household behavior. Second, this approach has the potential to allow for some generality as our specification of the investment process can be consistent with multiple models of the household. Other recent work derives parental endogenous behavior from explicit models of the household, including explicit representations of household preferences, decision making, beliefs, and constraints (see for example Bernal, 2008; Del Boca et al., 2014a,b; Cunha, 2013; Cunha et al., 2013; Doepke and Zilibotti, 2017; Agostinelli, 2019; Doepke et al., 2019; Agostinelli et al., 2020). The advantage of these latter approaches is that the counterfactual policy analysis incorporates well defined household responses to policy, see Del Boca et al. (2014b) for some discussion.

Given the investment function does not derive from an explicit model, we interpret the parameters in a more "reduced-form" way. The parameter $\alpha_{1,t}$ can be interpreted as reflecting whether parents "reinforce" existing skill stocks ($\alpha_{1,t} > 0$) or "compensate" for low skill stocks ($\alpha_{1,t} < 0$). The parameters $\alpha_{2,t}$ and $\alpha_{3,t}$ reflect the extent to which the mother's skills relate to the quantity and quality of her parental investment as in the case where more skilled mothers read to their children more or provide higher quality interactions. Finally, the parameter $\alpha_{4,t}$ reflects the influences that household resources have on the extent of parental investments, and it includes the combined effects of constraints the household faces (such as credit market constraints), as well as the household's preferences for investing scarce resources in children (see Caucutt et al., 2015).

Finally, to close the investment model, we assume that log family income ($\ln Y_t$) follows an AR(1) process:

$$\ln Y_{t+1} = \mu_Y + \rho_Y \ln Y_t + \eta_{Y,t} \tag{9}$$

where the innovation is $\eta_{Y,t}$ i.i.d. $\sim N(0, \sigma_Y^2)$ and is assumed independent of all latent variables. As specified above, initial family income Y_0 is allowed to be correlated with mother's and children's initial skills, and hence our model captures important correlations between household resources and the skills of parents and children.[17]

4.1.3 Skill Technology

The skill technology is specified according to the translog form introduced earlier in Equation 3:

$$\ln \theta_{t+1} = \ln A_t + \gamma_{1,t} \ln \theta_t + \gamma_{2,t} \ln I_t + \gamma_{3,t} \ln I_t \cdot \ln \theta_t + \eta_{\theta,t},$$

We assume that both the stock of skills θ_t and I_t are unobserved and measured with error. The production shock $\eta_{\theta,t}$, representing omitted inputs, is assumed distributed $\eta_{\theta,t}$ i.i.d. $\sim N(0, \sigma^2_{\theta,t})$. Our investment specification allows investment flows I_t to be endogenous to the current stock of child skills, mother's skills, and time varying household income, but are assumed independent of the contemporaneous skill shock $\eta_{\theta,t}$.[18] Further, we assume the contemporaneous latent stock of skills θ_t is independent of the current production shock, although correlated with past shocks given the dynamic technology specification.

4.1.4 Adult Outcome

In order to provide a more meaningful metric to evaluate policy interventions, we relate "adult" outcomes to the stock of children's skills in the final period of the child development process (period $T = 4$ or age 13-14). This follows the "anchoring" concept of Cunha et al. (2010). Each adult outcome Q is determined by

$$Q = \mu_Q + \alpha_Q \ln \theta_T + \eta_Q, \tag{10}$$

where η_Q is assumed independent of $\ln \theta_T$. We use years of schooling measured at age 23 and log earnings at age 29 as adult outcomes. Schooling is an attractive adult outcome to use because it explains a large fraction of adult earnings and consumption, is largely determined at an early point in adulthood, and, unlike realized labor market earnings, does not suffer from a censoring issue due to endogenous labor supply.

[17] Note that as with the assumption about the joint distribution of the initial conditions, our specific model for the evolution of household income is not strictly necessary for estimation of the main model primitives, in this case the investment equation. But to conduct counterfactual simulations we need to simulate household income paths for any given draw from the joint distribution of initial child skills, mother skills, and income.

[18] Identification relaxing this assumption would likely require some valid instrument for latent investment, perhaps derived from some randomized experimental intervention.

4.1.5 Measurement

We allow for all of the key variables to be measured with error. The full measurement system for latent investment, child cognitive skills, and mother's cognitive (MC) and non-cognitive (MN) skills is given by

$$
\begin{aligned}
Z_{I,t,m} =& \mu_{I,t,m} + \lambda_{I,t,m} \ln I_t + \epsilon_{I,t,m} \text{ for all } t, m \\
Z_{\theta,t,m} =& \mu_{\theta,t,m} + \lambda_{\theta,t,m} \ln \theta_t + \epsilon_{\theta,t,m} \text{ for all } t, m \\
Z_{MC,m} =& \mu_{MC,m} + \lambda_{MC,m} \ln \theta_{MC} + \epsilon_{MC,m} \text{ for all } m \\
Z_{MN,m} =& \mu_{MN,m} + \lambda_{MN,m} \ln \theta_{MN} + \epsilon_{MN,m} \text{ for all } m
\end{aligned}
$$

We assume measurement errors are independent of each other contemporaneously, independent of each other over-time, independent of latent variables, and independent of production and investment shocks. We do not make any assumption about the marginal distributions of the measurement errors, but use the primitive independence assumptions to form internally-valid instruments, as described below.

4.2 Estimation Algorithm

Our estimation approach directly follows our identification approach in treating the measurement parameters as nuisance parameters which can be computed sequentially along with the primitive parameters of the model generating the latent variables. The estimation algorithm is robust to parametric distributional assumptions on the marginal distributions of latent variables and measurement errors, as is commonly imposed in the prior empirical literature. Following the estimation of the initial conditions using standard techniques, we sequentially estimate for each age the investment and production functions, followed by the measurement parameters for the measures used for that age. The sequential algorithm we develop has the advantage of tractability because our estimator does not require the simulation of the full model; the primitives of the production technology and investment functions can be estimated directly from data without simulation. In addition, another advantage of our approach over a joint estimation approach is by breaking the estimator into steps, we make the identification assumptions as transparent as possible. Of course, the disadvantage of our approach is a potential loss of efficiency from not estimating the parameters jointly and exploiting "cross-step" restrictions.[19]

[19] It should be noted that to compute counterfactual simulations, we do simulate the full model forward from the estimated initial conditions, using the estimated model primitives. But simulating

Step 0 (Estimate Initial Conditions and Initial Measurement Parameters)

First, we estimate the measurement parameters at the initial period (age 5-6) for both children's and mother's skills. To estimate these measurement parameters, we use ratios of covariances and measurement means. We choose one measure each for children's cognitive skills, mother's cognitive skills, and mother's non-cognitive skills as the normalizing measure (which we label $m = 1$, without loss of generality) and normalize the factor loading for these measures to be 1: $\lambda_{\theta,0,1} = 1$, $\lambda_{MC,1} = 1$, $\lambda_{MN,1} = 1$. We estimate the remaining factor loadings using the average of the covariances between all of the remaining measures, where each factor loading is computed from

$$\lambda_{\theta,0,m} = \frac{Cov(Z_{\theta,0,m}, Z_{\theta,0,m'})}{Cov(Z_{\theta,0,1}, Z_{\theta,0,m'})} \ \forall m \neq m' \,,$$
$$\lambda_{\omega,m} = \frac{Cov(Z_{\omega,m}, Z_{\omega,m'})}{Cov(Z_{\omega,1}, Z_{\omega,m'})} \ \forall m \neq m' \text{ and } \forall\, \omega \in \{MC, MN\} \,.$$

Given the normalization that log skills are mean 0 in the initial period, we compute the initial measurement intercepts as

$$\mu_{\theta,0,m} = E(Z_{\theta,0,m}) \ \forall m \,,$$
$$\mu_{\omega,m} = E(Z_{\omega,m}) \ \forall m \text{ and } \forall\, \omega \in \{MC, MN\} \,.$$

With the factor loading estimates in hand, we then estimate the initial period variance-covariance matrix Σ_Ω using variances and covariances in measures of skills and family income (assumed measured without error). This step provides estimates of the initial joint distribution of children's skills, mother's skills, and family income. In this initial step, we also estimate the parameters of the income process (9) using a regression of log family on lagged log family income.

Finally, given the estimates of the measurement parameters for children and mother skills, we form the following "residual" measures:

$$\widetilde{Z}_{\theta,0,m} = \frac{Z_{\theta,0,m} - \mu_{\theta,0,m}}{\lambda_{\theta,0,m}} \ \forall m \,,$$
$$\widetilde{Z}_{\omega,m} = \frac{Z_{\omega,m} - \mu_{\omega,m}}{\lambda_{\omega,m}} \ \forall m \text{ and } \forall\, \omega \in \{MC, MN\} \,.$$

counterfactuals in this way does not require assuming anything about the marginal distribution of measurement errors or latent variables (beyond the initial period).

We are now ready to estimate the investment function for period $t = 0$, where the investment in this first period depends on the initial child's skills and household characteristics (mother's skills and family income).

Step 1 (Estimate Investment Function Parameters):

Following the errors-in-variables formulation described above, we substitute a "raw" measure for investment $Z_{I,0,m}$ and a "residual" measure for each of the latent skills ($\widetilde{Z}_{\theta,0,m}$, $\widetilde{Z}_{MC,m}$, $\widetilde{Z}_{MN,m}$) into the model of investment defined in terms of primitives (8):

$$\frac{Z_{I,0,m} - \mu_{I,0,m} - \epsilon_{I,0,m}}{\lambda_{I,0,m}} = \alpha_{1,0}(\widetilde{Z}_{\theta,0,m} - \widetilde{\epsilon}_{\theta,0,m}) + \alpha_{2,0}(\widetilde{Z}_{MC,m} - \widetilde{\epsilon}_{MC,m})$$

$$+\alpha_{3,0}(\widetilde{Z}_{MN,m} - \widetilde{\epsilon}_{MN,m}) + \alpha_{4,0} \ln Y_0 + \eta_{I,0}$$

Re-arranging, we have

$$\begin{aligned} Z_{I,0,m} &= \mu_{I,0,m} + \lambda_{I,0,m}\alpha_{1,0}\widetilde{Z}_{\theta,0,m} + \lambda_{I,0,m}\alpha_{2,0}\widetilde{Z}_{MC,m} + \lambda_{I,0,m}\alpha_{3,0}\widetilde{Z}_{MN,m} + \lambda_{I,0,m}\alpha_{4,0}\ln Y_0 \\ &+ \epsilon_{I,0,m} + \lambda_{I,0,m}(\eta_{I,0} - \widetilde{\epsilon}_{\theta,0,m} - \widetilde{\epsilon}_{MC,m} - \widetilde{\epsilon}_{MN,m}) \\ &= \beta_{0,0,m} + \beta_{1,0,m}\widetilde{Z}_{\theta,0,m} + \beta_{2,0,m}\widetilde{Z}_{MC,m} + \beta_{3,0,m}\widetilde{Z}_{MN,m} + \beta_{4,0,m}\ln Y_0 + \pi_{I,0,m} \qquad (11) \end{aligned}$$

where $\widetilde{\epsilon}_{\theta,0,m} = \frac{\epsilon_{\theta,0,m}}{\lambda_{\theta,0,m}}$, $\widetilde{\epsilon}_{MC,m} = \frac{\epsilon_{MC,m}}{\lambda_{MC,m}}$, $\widetilde{\epsilon}_{MN,m} = \frac{\epsilon_{MN,m}}{\lambda_{MN,m}}$, $\beta_{j,0,m} = \lambda_{I,0,m}\alpha_{j,0}$ for all j and

$$\pi_{I,0,m} = \epsilon_{I,0,m} + \lambda_{I,0,m}(\eta_{I,0} - \alpha_{1,0}\widetilde{\epsilon}_{\theta,0,m} - \alpha_{2,0}\widetilde{\epsilon}_{MC,m} - \alpha_{3,0}\widetilde{\epsilon}_{MN,m}).$$

Estimation of (11) by OLS would yield inconsistent estimates of the $\beta_{j,0,m}$ coefficients because the measures are correlated with their measurement errors (included in the residual term $\pi_{I,0,m}$). Here the structure of the model affords the researcher several possible strategies to consistently estimate the $\beta_{j,0,m}$ coefficients. We use an instrumental variable estimator with the vector of excluded instruments composed of alternative measures of skills: [$Z_{\theta,0,m'}$, $Z_{MC,m'}$, $Z_{NC,m'}$]. Under our previous assumptions about the measurement errors, these instruments are valid because the alternative measures are uncorrelated with all of the components of $\pi_{0,m}$. Using this IV strategy, we obtain consistent estimators for the $\beta_{j,t,m}$ coefficients. The primitive parameters of the investment function are then recovered from

$$\alpha_{j,0} = \frac{\beta_{j,0,m}}{\sum_{j=1}^{4} \beta_{j,0,m}} \ \ \forall j \in \{1, \ldots, 4\}$$

Step 2 (Compute Measurement Parameters for Latent Investment):

After estimating the primitive parameters of the investment function, we recover the scale and location for the investment equation without further assumptions on the measurement equation parameters. The intercept and factor loading for the investment measure are given by

$$\mu_{I,0,m} = \beta_{0,0,m} \,,$$

and

$$\lambda_{I,0,m} = \sum_{j=1}^{4} \beta_{j,0,m} \,.$$

With these consistent estimators for the measurement parameters, we form the "residual" measures for investment in period $t = 0$:

$$\widetilde{Z}_{I,0,m} = \frac{Z_{I,0,m} - \mu_{I,0,m}}{\lambda_{I,0,m}} \equiv \ln I_0 + \widetilde{\epsilon}_{I,0,m} \,,$$

where $\widetilde{\epsilon}_{I,0,m} = \frac{\epsilon_{I,0,m}}{\lambda_{I,0,m}}$.

Step 3 (Estimate Skill Production Technology)

We assume we have available at least one child skill measure which is age-invariant. Label the age-invariant measure to be measure m, and for this measure we have $\mu_{\theta,t,m} = \mu_{\theta,0,m}$ and $\lambda_{\theta,t,m} = \lambda_{\theta,0,m}$ for all t.

Substituting the residual measures into the production technology (3), we have

$$\frac{Z_{\theta,1,m} - \mu_{\theta,1,m} - \epsilon_{\theta,1,m}}{\lambda_{\theta,1,m}} = \ln A_0 + \gamma_{1,0}(\widetilde{Z}_{\theta,0,m} - \widetilde{\epsilon}_{\theta,0,m}) + \gamma_{2,0}(\widetilde{Z}_{I,0,m} - \widetilde{\epsilon}_{I,0,m})$$
$$+ \gamma_{3,0}(\widetilde{Z}_{\theta,0,m} - \widetilde{\epsilon}_{\theta,0,m})(\widetilde{Z}_{I,0,m} - \widetilde{\epsilon}_{I,0,m}) + \eta_{\theta,0} \,.$$

With some algebra, we can re-write this as:

$$Z_{\theta,1,m} = \delta_{0,0,m} + \delta_{1,0,m}\widetilde{Z}_{\theta,0,m} + \delta_{2,0,m}\widetilde{Z}_{I,0,m} + \delta_{3,0,m}\widetilde{Z}_{\theta,0,m} \cdot \widetilde{Z}_{I,0,m} + \pi_{\theta,0,m} \,, \qquad (12)$$

where the new error term $\pi_{\theta,0,m}$ is:

$$\pi_{\theta,0,m} = \epsilon_{\theta,1,m} + \lambda_{\theta,1,m}[\eta_{\theta,0} - \gamma_{1,0}\widetilde{\epsilon}_{\theta,0,m} - \gamma_{2,0}\widetilde{\epsilon}_{I,0,m} - \gamma_{3,0}(\widetilde{Z}_{\theta,0,m}\widetilde{\epsilon}_{I,0,m} + \widetilde{Z}_{I,0,m}\widetilde{\epsilon}_{\theta,0,m} - \widetilde{\epsilon}_{\theta,0,m}\widetilde{\epsilon}_{I,0,m})] \,.$$

The rest of the reduced-form parameters (δs) map into the structural parameters and measurement parameters in the following way:

$$\delta_{0,0,m} = \mu_{\theta,1,m} + \lambda_{\theta,1,m} \cdot \ln A_0$$

$$\delta_{j,0,m} = \lambda_{\theta,1,m}\gamma_{j,0} \text{ for any } j \in \{1,2,3\}.$$

As with the investment function, estimation of (12) using OLS would lead to an inconsistent estimator. The vector of excluded instruments is composed of alternative measures of skills and investment: $[Z_{\theta,0,m'}\ , Z_{I,0,m'}\ , Z_{\theta,0,m'} \cdot Z_{I,0,m'}]$. Under our measurement assumptions, these instruments are valid because the alternative measures are uncorrelated with all of the components of $\pi_{\theta,0,m}$.[20] With a consistent estimator of the δs in hand, we can then recover the structural parameters:

$$\ln A_0 = \frac{\delta_{0,0,m} - \mu_{\theta,1,m}}{\lambda_{\theta,1,m}}$$

$$\gamma_{j,0} = \frac{\delta_{j,0,m}}{\lambda_{\theta,1,m}} \ \forall \ j \in \{1,2,3\} \ ,$$

where both measurement parameters $\mu_{\theta,1,m}$ and $\lambda_{\theta,1,m}$ are already identified under the age-invariance assumption.

Step 4 (Estimate Variance of Investment and Production Function Shocks):

The variances of both the investment shocks ($\sigma^2_{I,0}$) and of the production function shocks ($\sigma^2_{\theta,0}$) remain to be estimated. In order to estimate $\sigma^2_{I,0}$, we use the covariance between the residual term ($\pi_{I,0,m}$) in (11), and an alternative residual measure of investment $\widetilde{Z}_{I,0,m'} = \ln I_0 + \widetilde{\epsilon}_{I,0,m'}$ as follows:

$$Cov\left(\frac{\pi_{I,0,m}}{\lambda_{I,0,m}}, \widetilde{Z}_{I,0,m'}\right) = V(\eta_{I,0}) = \sigma^2_{I,0} \, .$$

To compute the residual measure $\widetilde{Z}_{I,0,m}$, we need to compute the measurement parameters for this measure. We do this by following the procedure explained in

[20] Perhaps the less obvious terms are terms such as this $E(\widetilde{Z}_{\theta,0,m}\epsilon_{I,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'})$. Under the assumption of independence of the errors, we have

$$E(\widetilde{Z}_{\theta,0,m}\epsilon_{I,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'}) = E(\widetilde{Z}_{\theta,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'})E(\epsilon_{I,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'})$$

given $\epsilon_{I,0,m}$ is independent of $\widetilde{Z}_{\theta,0,m}$. Given the independence assumption, the latter term is $E(\epsilon_{I,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'}) = E(\epsilon_{I,0,m}) = 0$. Therefore, $E(\widetilde{Z}_{\theta,0,m}\epsilon_{I,0,m}|Z_{\theta,0,m'} \cdot Z_{I,0,m'}) = 0$.

the Steps 1 and 2 above, with the alternative measure $Z_{I,0,m'}$. The variance of the production shock is estimated in the same way using an alternative measure of children's skills in period $t = 1$:

$$Cov\left(\frac{\pi_{\theta,1,m}}{\lambda_{\theta,1,m}}, \widetilde{Z}_{\theta,1,m'}\right) = V(\eta_{\theta,0}) = \sigma^2_{\theta,0}\,.$$

Remaining Steps

We repeat Steps 1-4 for the remaining periods until the final period of child development T. This algorithm produces estimates of the parameters for all child ages.

Finally, after we have computed the full path of primitive parameters for the investment and production functions, we are able to estimate the adult outcome process (10). We estimate equations for both years of education at age 23 and log earnings at age 29. We use the same IV method as before to solve the measurement error issue. Substituting the measures for skills at age 13-14 ($T = 4$) in equation (10), we have:

$$Q = \mu_Q + \alpha_Q \widetilde{Z}_{\theta,4,m} + (\eta_Q - \alpha_Q \widetilde{\epsilon}_{\theta,4,m}) \tag{13}$$

where we use a second measure for skills at age 13-14 as an instrumental variable to identify α_Q.

Standard Errors

To account for the sources of estimation uncertainty among different steps and the sample design of the data, we adopt a bootstrap algorithm. Because of the original sampling design of the data, we implement a block bootstrap algorithm that samples all the children of each family to account for the intra-family correlation, and all of the observations at each age for each child to account for the intra-child correlation. We generate a distribution of estimated parameters by re-estimating the model, including all estimation steps in sequence, for each of the random samples of families. This distribution is used to compute the standard errors and confidence intervals.

4.3 Data

We estimate the model using information about children and their families obtained from the National Longitudinal Study of Youth 1979 (NLSY). Descriptive statistics for the sample and additional data construction details are left for the Appendix.

The NLSY dataset is constructed by matching female respondents of the original dataset with their children who were part of the Children and Young Adults surveys, from 1986 to 2012. The dataset provides observations of the first period of the model (age 5-6) through adulthood. The total number of children in our sample is 11,509.

The NLSY dataset contains multiple measures of children's skills, mother's skills, and parental investments. The complete set of measures, their ranges and descriptive statistics for our sample are included in the Appendix. For children's skills we rely on different sub-scales of the Peabody Individual Achievement Test (PIAT) in Mathematics, Reading and Recognition, and the Peabody Picture Vocabulary Test (PPVT). Finally, we use information for children when they become young adults to link the children's skills into a more meaningful metric to evaluate policy intervention: we use children's highest grade completed at age 23 or older and their earnings at age 29. The information about the educational attainment is measured as the highest grade completed as of date of last interview. We considered schooling information only for those young adults who were at least 23 years old or older in the last 2012 interview. Age 29 earnings is in real 2012 dollars.

For mother's cognitive skills we use sub-scales of the Armed Services Vocational Aptitude Battery (ASVAB), and for mother's non-cognitive skills we use the Rotter and Rosenberg indexes. For parental investments, we use the various HOME score measures from direct observation and interview with the mother. Family income includes all sources of income for the parents, including mother's and father's labor income, and any sources of non-labor income.

Finally, we rely on PIAT-Math scores for our age-invariant measure, from age 5 to 14. As discussed in Ollendick and Cerny (2013), the PIAT was designed to provide a reliable measurement of the cognitive developmental level of children. The authors discuss that the validity of the PIAT has been measured by observed developmental changes, correlations with other tests, and factor analysis. Dunn and Markwardt (1970) reviewed the psychometric properties of the PIAT measures. Developmental changes in children has been shown to be strongly associated with an increase in scores with age or grade.

In the Appendix, we describe the testing procedure for these assessments (taken from the NLSY codebooks). The PIAT-Math consists of 84 items of increasing difficulty. For our sample, the mean score increases from about 12 items correct for children aged 5-6 to 54 items correct for children aged 13-14. Less than 1 percent of the sample scored at the maximum value of the test (all 84 items correct) or at the minimum value (0 items correct). In addition, there is substantial variation in scores at each age. As we note above, the absence of a binding floor or ceiling in the measure is not proof of age-invariance, but does suggest the measure is not merely appropriate for a subset of children.

5 Results

In this Section we discuss our parameter estimates and simulate the estimated model to describe the development of children's skills. We conclude this Section with a series of policy counterfactual experiments using the estimated model. These exercises provide a metric to interpret the estimates with respect to adult outcomes, schooling and earnings.

5.1 Parameter Estimates

5.1.1 Initial Conditions

Table B-1 reports estimates of the initial conditions variance-covariance matrix Σ_Ω and the associated correlation matrix. We normalize children's cognitive skills to the PIAT-Mathematics test, mother's cognitive skills to the ASVAB2 (Arithmetics reasoning) and mother's non-cognitive skills to the Self-Esteem 1 (Rosenberg Self-Esteem: "I am a person of worth") measure. The variances and covariances of the latent skills, and the investment and production function parameters, are interpreted relative to these normalizations. As expected, we estimate that children's skills, mother's cognitive and non-cognitive skills, and family income are all highly positively correlated. For space considerations, estimates of the dynamic family income process can be found in the Appendix.

5.1.2 Investment Function

Table B-2 reports the estimates of the investment function specified in Section 4.1.2. At ages 5-6, we find that investment is increasing in children's skills, mother's skills, and family income. Because of the log-log form of the investment equation, we can interpret parameter estimates as elasticities. However, we note that the interpretation of these estimates still depends on the particular normalizations and log-linear form of the measurement equations. The parameter estimate of 0.230 on the log children's skills variable indicates that a 1 percent increase in children's skills raises investment by 0.23 percent, an inelastic response. The positive coefficient suggests that parents are "reinforcing" existing skills with further investments: children with higher skills are receiving even more investment than children with lower skills. Mother's cognitive and non-cognitive skills also increase investment at ages 5-6, with non-cognitive skills of the mother estimated to have a substantially higher elasticity than cognitive skills. These coefficients indicate that mothers with higher skills are providing higher investments in children. Turning to the importance of income to parental investments, we find that a 1 percent increase in family income raises investment by 0.34 percent. The response of investment with respect to mother's skills and family income reflects the combination of parental preferences and household constraints, which we cannot unfortunately separately distinguish using this reduced-form model of investment. Given that positive correlation between mother's initial skills, child's initial skills, and household income, these estimates of the investment function indicate that endogenous investment increases inequality in children's skills. The estimated variance of the investment shock reveals how much of the remaining variation in parental investments remains unexplained by this model, such as investments from schools, peers, and the child herself.

Comparing parameter estimates of the investment function over the development period reveals that the influence of the child's prior skills on investments becomes much smaller at later ages, indicating that parental investments become less reinforcing of existing skill stocks at older ages. As the child develops, we find that the mother's non-cognitive skills become the dominant influence on investment. However, while the importance of family income falls somewhat from an elasticity of 0.34 at age 5-6 to 0.275 at age 11-12, income is still a significant and positive factor for parental investment even at later ages.

5.1.3 Production Function

Table 2 reports the parameter estimates for the technology of skill formation, as described in Section 2.3. At all ages, we find that skills are "self-productive" (next period's skills are increasing in existing skill stocks) and that skills are positively increasing in investment. For age 5-6 skill production, we estimate a significant negative complementarity between the stock of a child's skills and parental investments (the interaction term $\ln \theta_t \ln I_t$). This result highlights the importance of departing from the Cobb-Douglas/CES specifications.

The elasticities of skill production with respect to investment are heterogeneous, and we graph the skill elasticity for the age 5-6 production function in Figure 1 with respect to the existing stock of children's skill. The estimated negative coefficient on the interaction term indicates that the elasticity of skill production with respect to investment is decreasing in the child's current skill level. For low skill children, the elasticity approaches 1.5, indicating that a 1 percent increase in investment increases next period's skills by 1.5 percent. For already high skill children, the elasticity approaches 0.2, indicating that a 1 percent increase in investment raises future skills by only 0.2 percent.

The heterogeneous investment elasticities suggest that interventions would have the largest effect on skill disadvantage children. This result stands in contrast to the estimates from previous works in the literature, which were based on CES (or linear) technologies. These technology specifications restrict the heterogeneity of the investment productivity by assuming that the marginal productivity of investment must be *increasing* (or constant) with respect to a child's skills. Note also that unlike the CES (constant returns to scale) case, our unrestricted model allows investment elasticities to be larger than 1. Indeed, our estimates suggest that, at least for some children, skill production is relatively highly elastic with respect to investment.

The high TFP estimate for age 5-6 and the increasing returns to scale indicate that existing skills and investments at this initial age are very productive relative to later ages; early periods are more sensitive. These estimates of high returns to early investment will underlie the policy experiment results we discuss next. As children age, Table 2 indicates that skills and investments become generally less productive and skills less "malleable." We graph the estimated TFP at each age in Figure 2. Our estimate of TFP at age 11-12 falls to 1/6 the level at age 5-6, indicating a dramatic slowdown in the productivity of existing skills and investments in producing new skills. This feature of the technology is largely consistent with the evidence that cognitive skills are difficult to change as children reach age 10.[21]

5.1.4 Adult Outcomes

Table B-3 presents our estimates of the completed schooling outcome equation and log earnings equation. We estimate that a 1 unit increase in children's log skills at age 13-14 leads to an increase of 0.15 years of school and a 0.041 increase in log earnings at age 29. Below, we use these estimates to "anchor" our policy estimates to a meaningful adult outcome metric.

5.2 Estimated Child Development Path

We analyze the quantitative implications of the estimated model by simulating the dynamic model. Simulation of the model proceeds by drawing 10,000 children from the estimated initial conditions distribution, and, for each child, forward simulating the path of income, investments, children's skills, and adult outcomes.

[21] In the Appendix, we report estimates of the technology using alternative sample definitions and adding additional covariates. The results are qualitatively similar.

Figure 3 shows the estimated development path of mean log latent cognitive skills. Figures 4 and 5 show the dynamics in the distribution of latent skills. And, Figure 6 provides the estimated dynamics in the distribution of latent investment.

Perhaps not surprisingly, we find that children's mean latent skills grow substantially over this development period, from age 5 to 14, with the most rapid growth at early ages and growth slowing somewhat in the later period. As discussed above, key to the identification of the non-stationary change in children's skills is the use of age-invariant measures of children's skills. In addition to growth in mean skills, we estimate that the latent distribution of cognitive skills becomes more dispersed as children age (Figure 4). Inequality rises substantially as there are different rates of skill growth for children at different percentiles of the initial skill distribution. Figure 5 shows that skills for high skill children at the 90^{th} percentile grow 20% from age 5-6 to age 9-10 and grow 9% during the rest of the childhood. For low initial skill children at the 5^{th} percentile, growth is slower, with a 6 % growth rate from age 5-6 to age 7-8 and a 3 % growth rate from age 11-12 to age 13-14.

5.3 Policy Experiments

In this sub-section, we use the estimated model to analyze the effect of income transfers on childhood skill development and adult outcomes. We argue that this experiment provides a meaningful metric to understand the magnitude of the parameter estimates, as well as allowing us to study the possible trade-offs of designing childhood policies.

5.3.1 Effects on Final Skills

We first consider a simple exercise designed to assess the optimal timing of the income transfer. In Figure 7 we show the average change in the latent children's log skills at age 13-14 by the different timing (age) of income transfer: $E[\ln \theta'_T(a) - \ln \theta_T]$, where $\theta'_T(a)$ is the level of skill at age $t = T$ (age 13-14) with an income transfer of $10,000 dollars (in 2012 $) provided to the family at age $t = a$, and θ_T is level of skill at age 13-14 in the baseline estimates (no income transfer). The transfer is a one-time transfer, and we do not allow households to save past transfers or borrow from future transfers. The figure shows that a $10,000 transfer given at age 5-6 increases the average stock of age 13-14 skills by approximately 16 percent. Providing the same transfer later has a smaller average effect. Providing a $10,000 transfer at age 11-12 would increase the average skill stock at age 13-14 by less than 3 percent. This result reflects the high productivity of investment in the early periods and the high level of productivity of existing stocks of skill in producing future skills.

5.3.2 Effects on Completed Schooling

Figure 8 displays the results of the same set of policy experiments as in Figure 7 but using completed schooling at age 23 as the outcome. In this Figure, we plot $E(S^{'}(a) - S)$, where $S^{'}(a)$ is the number of months of completed schooling for an income transfer of \$10,000 given at age a, and S is the number of months of completed schooling in the baseline model (no income transfer). We find that a \$10,000 transfer given at age 5-6 would increase the number of average months of completed schooling by 0.3 months. Providing the same transfer at a later period would increase completed schooling by only 0.05 months.

5.3.3 Heterogeneous Treatment Effects

The previous results showed the average effect of policies providing transfers at different stages of the development process. Our modeling framework allows for potentially important sources of heterogeneity by the child's initial skills, mother's skills, and initial family income levels; all of which could affect the individual level treatment effect. The model estimates allow us to directly estimate this heterogeneity in the policy treatment effects.

Figure 9 plots the heterogeneous effect of the income transfer policy at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income. This figure also plots the average treatment effect (ATE), the average effect over the income distribution (the same effect as reported above). While the ATE is about 0.3 months, the effect varies considerably depending on the child's initial level of income. For the children from poor households in the lowest income percentiles, the effect of the income transfer is to increase completed schooling by 1 month or more, and for the children from the richest households, the effect is near 0. The large heterogeneous effects by family income stem from the estimated importance of family income in producing child investments and the estimated positive correlation of income with maternal skills and the child's initial skills. This heterogeneity in the effects by income mirrors the heterogeneity in income effects found in previous papers using alternative sources of identification (see Dahl and Lochner, 2012; Loken et al., 2012; Agostinelli and Sorrenti, 2018). Using the varied effects of the Norwegian oil boom to instrument for family income, Loken et al. (2012) report estimates on completed schooling which are smaller in magnitude than those reported here, but similar qualitatively in finding that the effects are substantially larger for low income Norwegian families.

Figure 10 plots the heterogeneous effect of the same policy by the level of the child's initial (age 5-6) skill. The ATE plotted in this Figure is the same as in the previous figure as it is simply the effect averaged over the initial skill distribution. In this Figure, we also see evidence of heterogeneous treatment effects, with low initial skill children benefiting more (up to about 0.8 months of additional schooling) from the policy intervention than high initial skill children (near 0 effect). But the importance of heterogeneity by initial skill is substantially smaller than by family income. This suggests that it is better to target the policy to low income households than low skill households, but of course it cannot be worse to target based on both criteria.

5.4 Quantifying the Importance of Measurement Error

Our results presented thus far have been for the model with measurement error corrections, estimated using what we argue are skill measures that are age-invariant. We next briefly discuss how the estimates of the primitive production technology would differ if we ignore the measurement error issues. Note that there is no clear theoretical prediction about the sign of the measurement error bias, given that our models are dynamic, non-linear, and consist of inter-related multiple equations. Our analysis allows us to quantify the importance of measurement error, using policy predictions on adult schooling as a meaningful metric for comparison.

Table 3 presents the estimated policy effects for two specifications, measurement error corrected and uncorrected. In Panel A of Table 3, we present the average treatment effects (ATE) on adult schooling of the $10,000 income transfer at various ages for both estimators. The first row shows the estimates that account for measurement error: we estimate that the income transfer at age 5-6 would increase average schooling by 0.3 additional months. The second row shows the estimated ATE if we do not correct for measurement error. Using these uncorrected estimates, we estimate policy effects that are almost 5 times smaller than the measurement error corrected estimates, a substantial reduction in the estimated effect of an income transfer.

Panel B of Table 3 repeats the analysis but focusing on the heterogeneity in the treatment effect at different parts of the family income distribution. Similar conclusions are evident here: ignoring measurement error leads to a substantial reduction in the estimated policy effects of the income transfer. For example, we see that ignoring measurement error would lead to estimated policy effects on low income families at the 10th percentile that are five times lower than the estimated effects once we account for measurement error.

5.5 Cost-Benefit Analysis

We have thus far shown that the estimated model implies that a policy intervention of providing income transfers to families would produce modest but positive gains in children's skills, with larger effects for poorer households. Would these gains be justified given the cost? We next present a simple cost-benefit analysis that focuses on an income transfer policy to the needy families in our sample (at the 10th percentile of initial family income).

Table 4 shows the effects of the income transfer policy, by children's age, on the present value of earnings. In our cost-benefit analysis we calculate the net benefits of the income transfer at age 5-6 by family income in terms of the net present value of future earnings. The cost of the policy includes the direct transfer, as well as the cost of additional schooling. The benefit of this policy is the comparison between the present value of worker's earnings with and without that income transfer during childhood. In other words, we compute the counterfactual present value of earnings if the worker's family had received the income transfer when the worker was a child under different scenarios about the baseline adult earnings. The effect of the family income transfer to the growth in children earnings are computed using estimates in Table B-3 under the assumption that the change in the growth rate due to the policy intervention is constant over the life-cycle.[22]

Table 4 suggests that, considering both the cost of the income transfer and the cost of additional education, the net benefit of the policy is positive for earlier ages and it declines throughout childhood. The largest policy net benefit is when children are 5-6 years old, and it turns negative if the income transfer is implemented too late, when the child is 11-12 years old. The present value for the policy intervention at age 5-6 is slightly more than $29,000 and the net benefit is around $18,000.

6 Conclusion

This paper develops new identification results based on measurement restrictions on skill measures. These empirically grounded assumptions are sufficient for identification of general skill development technologies, including features that were not previously considered, total factor productivity and returns to scale. Based on our identification results, we develop a robust estimator using a sequential estimating algorithm. Our estimator does not require strong assumptions about the marginal distribution of measurement errors or the latent factors.

[22]We assume that the baseline adult earnings are constant over the life cycle. Allowing for earnings growth would generate higher benefits of the policy.

We estimate the skill production process using data for the United States and a flexible parametric model of skill development allowing free skill production complementarity between a child's stock of skills and parental investments. Our parameter estimates reveal that investments are more productive at early ages and in particular for disadvantaged children. Our findings of a positive return to income transfers at early ages, especially for poorer households, is largely consistent with prior evidence of a positive effect of income on a number of child outcomes (see Dahl and Lochner, 2012; Loken et al., 2012; Agostinelli and Sorrenti, 2018) using different sources of identification. Our results suggest that family income is a better "target" than initial children's skills for children's skills, with positive net returns when the policy targets low-income families. Lastly, our finding that the estimated policy effects are attenuated by measurement error bias demonstrates the critical importance of correcting estimates for measurement error.

References

Karun Adusumilli and Taisuke Otsu. Nonparametric instrumental regression with errors in variables. *Econometric Theory*, 34(6):1256–1280, 2018.

Francesco Agostinelli. Investing in children's skills: An equilibrium analysis of social interactions and parental investments. *Working Paper*, 2019.

Francesco Agostinelli and Giuseppe Sorrenti. Money versus time: family income, maternal labor supply, and child development. *Working Paper 2018-017, HCEO, University of Chicago*, 2018.

Francesco Agostinelli and Matthew Wiswall. Identification of dynamic latent factor models: The implications of re-normalization in a model of child development. *Working Paper*, 2016.

Francesco Agostinelli, Matthew Wiswall, and Morteza Saharkhiz. Home and school in the development of children. *NBER Working Paper*, (26037), 2019.

Francesco Agostinelli, Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti. It takes a village: The economics of parenting with neighborhood and peer effects. *Working Paper*, 2020.

Orazio Attanasio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. Estimating the production function for human capital: Results from a randomized control trial in colombia. *NBER Working Paper*, (21740), 2015a.

Orazio Attanasio, Costas Meghir, and Emily Nix. Human capital development and parental investment in india. *NBER Working Paper*, (20965), 2015b.

Orazio Attanasio, Flávio Cunha, and Pamela Jervis. Subjective parental beliefs: Their measurement and role. *NBER Working Paper*, 2019a.

Orazio Attanasio, Costas Meghir, and Emily Nix. Human capital development and parental investment in india. Technical report, National Bureau of Economic Research, 2019b.

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

Dale Ballou. Test scaling and value-added measurement. *Education Finance and Policy*, 4(4):351–83, 2009.

Raquel Bernal. The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4):1173–1209, 2008.

Timothy N. Bond and Kevin Lang. 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, 2013.

Timothy N. Bond and Kevin Lang. The black-white education-scaled test-score gap in grades k-7. *Journal of Human Resources*, Fall(53):891–917, 2018.

Elizabeth M. Caucutt, Lance Lochner, and Youngmin Park. Correlation, consumption, confusion, or constraints: Why do poor children perform so poorly? *NBER Working Papers*, (21023), 2015.

Gary Chamberlain. Education, income, and ability revisited. *Journal of Econometrics*, pages 241–257, 1977.

Raj Chetty, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. How does your kindergarten classroom affect your earnings? evidence from project star *. *The Quarterly Journal of Economics*, 126(4):1593–1660, 2011.

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632, September 2014.

Flavio Cunha. Investments in children when markets are incomplete. *Working Paper*, 2013.

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007.

Flavio Cunha and James J. Heckman. Formulating, identifying, and estimating the technology for the formation of skills. *Journal of Human Resources*, 43(4):738–782, 2008.

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

Flavio Cunha, Eric Nielsen, and Benjamin Williams. The econometrics of early childhood human capital and investments. *The Annual Review of Economics*, 13: 487–513, 2021.

Flávio Cunha, Irma Elo, and Jennifer Culhane. Eliciting maternal expectations about the technology of cognitive skill formation. *NBER working paper*, (19144), 2013.

Gordon B. Dahl and Lance Lochner. The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5):1927–1956, 2012.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household choices and child development. *Review of Economic Studies*, 81(1):137–185, 2014a.

Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Transfers to households with children and child development. *Working Paper*, 2014b.

Emilia Del Bono, Josh Kinsler, and Ronni Pavan. A note on the importance of normalizations in dynamic latent factor models of skill formation. *Working Paper*, 2020.

Matthias Doepke and Fabrizio Zilibotti. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371, 2017.

Matthias Doepke, Giuseppe Sorrenti, and Fabrizio Zilibotti. The economics of parenting. *Annual Review of Economics*, 11(1):55–84, 2019.

L Dunn and L Dunn. The peabody picture vocabulary test-revised. *Circle Pines, MN: American Guidance Service*, 1981.

L. M. Dunn and Jr. Markwardt, F. C. *Peabody Individual Achievement Test*. Circle Pines, Minn.: American Guidance Service, 1970.

Joachim Freyberger. Normalizations and misspecification in skill formation models. *Working Paper*, 2020.

J. Hausman, H. Ichimura, W. Newey, and J Powell. Measurement errors in polynomial regression models. *Journal of Econometrics*, 50:271–295, 1991.

James J. Heckman and Stefano Mosso. The economics of human development and social mobility. *NBER Working Paper*, (19925), 2014.

M. J. Kolen and R. L. Brennan. *Test equating, scaling, and linking: Methods and practices*. Springer, New York, NY, 2014.

Michael Lewis and Harry McGurk. Evaluation of infant intelligence. *Science*, 178 (4066):1174–7, 1972.

Katrine V. Loken, Magne Mogstad, and Matthew Wiswall. What linear estimators miss: Re-examining the effects of family income on child outcomes. *American Economic Journal: Applied Economics*, 4(2):1–35, 2012.

H. Thomas Ollendick and A. Jerome Cerny. *Clinical Behavior Therapy with Children*. Springer Science and Business Media, 2013.

Susanne M. Schennach. Estimation of nonlinear models with measurement error. *Econometrica*, (72):33–75, 2004.

Benjamin D. Williams. Identification of a nonseparable model under endogeneity using binary proxies for unobserved heterogeneity. *Quantitative Economics*, 10(2): 527–563, 2019.

Table 1: Sample Descriptive Statistics

	Mean	SD
N Obs	19070	
N of Mothers	3199	
N of Children	4941	
% Male Children	51.32	
% Female Children	48.68	
% Hispanic Children	21.44	
% Black Children	30.44	
% Other Races	48.12	
Mom Education	12.59	2.63
Family Income	61657.88	47527.85
Children Final Years of Education	13.30	2.36

Notes: This table shows the main descriptive statistics of the CNLSY79 sample we use to estimate the model. Children's Completed Education is the child's completed years of education at age 23. The variable "other races" represents all children which are not black neither Hispanic (i.e. it includes white, non-Hispanic children). Income is in $2012 USD.

Table 2: Estimates for the Skill Technology

Model	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	1.966	1.086	0.897	1.065
	(0.153)	(0.035)	(0.027)	(0.029)
	[1.71,2.21]	[1.03,1.15]	[0.85,0.93]	[1.02,1.11]
Log Investments	0.799	0.695	0.713	0.252
	(0.261)	(0.337)	(0.403)	(0.538)
	[0.43,1.22]	[0.16,1.21]	[-0.03,1.24]	[-0.53,1.16]
Log Skills $\times$	-0.105	-0.005	-0.003	0.003
Log Investments	(0.066)	(0.019)	(0.013)	(0.010)
	[-0.21,-0.03]	[-0.04,0.03]	[-0.02,0.02]	[-0.02,0.02]
Standard Deviation	5.612	4.519	3.585	4.019
Shocks	(0.173)	(0.184)	(0.180)	(0.246)
	[5.38,5.93]	[4.28,4.88]	[3.27,3.87]	[3.71,4.43]
Log TFP	13.067	14.747	11.881	2.927
	(0.294)	(0.365)	(0.538)	(0.952)
	[12.67,13.60]	[14.24,15.47]	[11.20,12.93]	[1.40,4.48]

Notes: This table shows the measurement error corrected estimates for the technology of skill formation. Each column shows the coefficients of the technology of skill formation at the given age. The dependent variable is log skills in the next period $t+1$, and the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

Table 3: Estimated Policy Effects and Measurement Error

Panel A: ATE by Age of Income Transfer

Model	Age of Income Transfer ($ 10,000) Age 5-6	Age 7-8	Age 9-10	Age 11-12
Measur Error Corrected	0.300 [0.170,0.430]	0.172 [0.044,0.299]	0.129 [-0.001,0.258]	0.050 [-0.058,0.157]
Not Corrected for Measur Error	0.062 [0.042,0.082]	0.029 [0.014,0.044]	0.022 [0.007,0.036]	0.013 [-0.000,0.027]

Panel B: ATE at age 5-6 by Family Income

	Low Income Families (10th Income Percentile)	High Income Families (90th Income Percentile)
Measur Error Corrected	0.642 [0.369,0.915]	0.069 [0.034,0.105]
Not Corrected for Measur Error	0.126 [0.085,0.166]	0.018 [0.012,0.023]

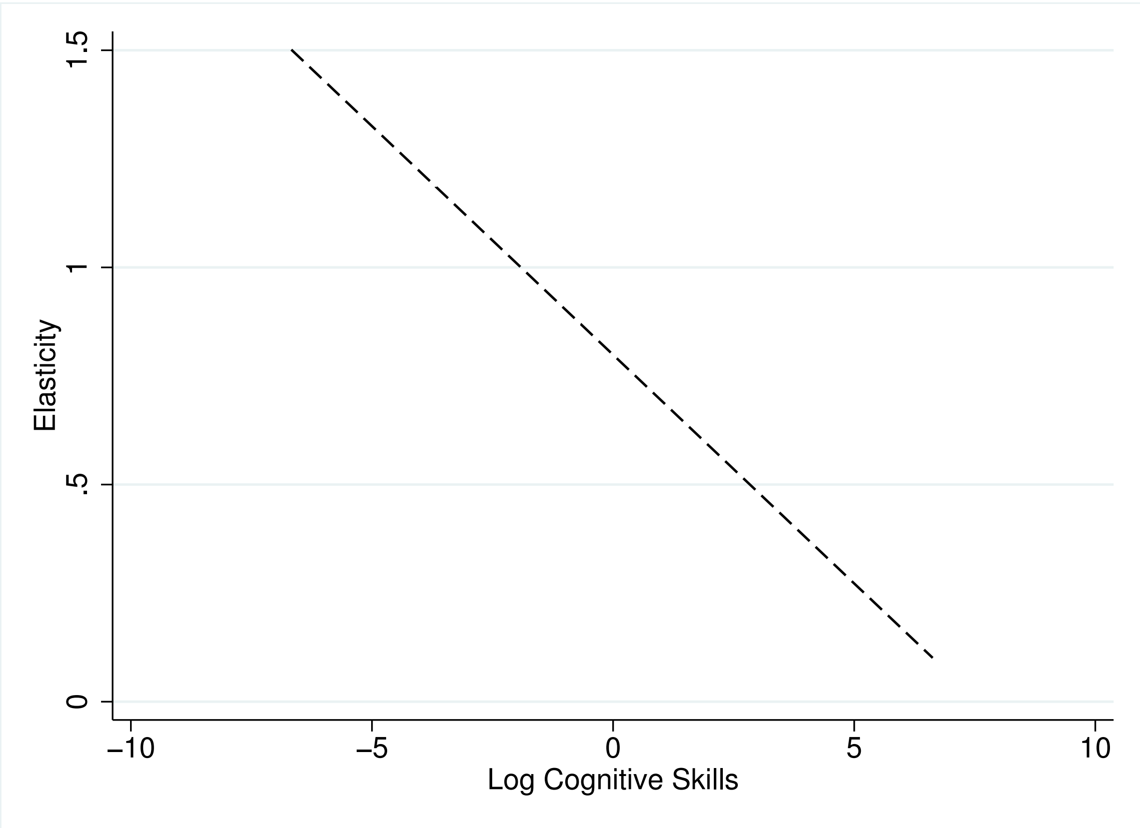
Notes: Panel A shows the average treatment effects (ATEs) on additional months of completed education by age of policy intervention ($10,000 income transfer) with and without correcting for measurement error. Panel B shows the ATEs by family income percentiles (10th vs 90th percentiles).

Table 4: Average Effect of an Income Transfer by Age of Transfer (Outcome: PV of Earnings)

Age of Intervention	Benefit on PV Earnings ($)	Direct Cost (Income Transfer) ($)	Cost of Education ($)	Net Benefit ($)
5	29126.81	10000	749.58	18377.23
7	15598.78	10000	403.65	5195.14
9	11541.43	10000	299.12	1242.31
11	4365	10000	113.43	-5748.43

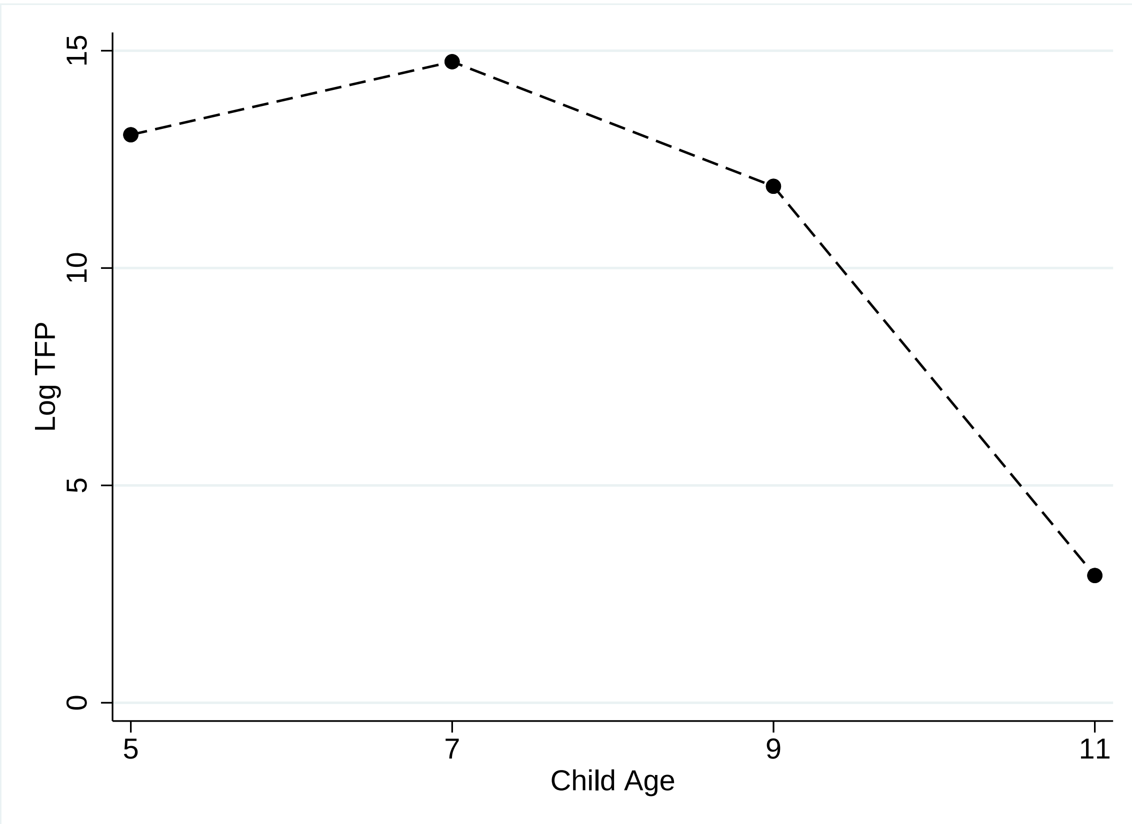
Notes: This table shows the benefit-cost analysis for a $10,000 dollars transfer to family. The benefit on the PV of earnings is the difference between the present value of earnings with and without that transfer when worker was age 5-6. The effect of family income transfer on earnings growth is computed adjusting for the increased earning growth implied by estimates in Table B-3. The cost of that policy takes into account both the direct transfer and the discounted cost of additional education that the policy induces. We use a yearly cost of school of 10,000 dollars per year. We discount future earnings with an interest rate of 2%.

Figure 1: Estimates of Skill Production Elasticity with Respect to Investment at Age 5-6



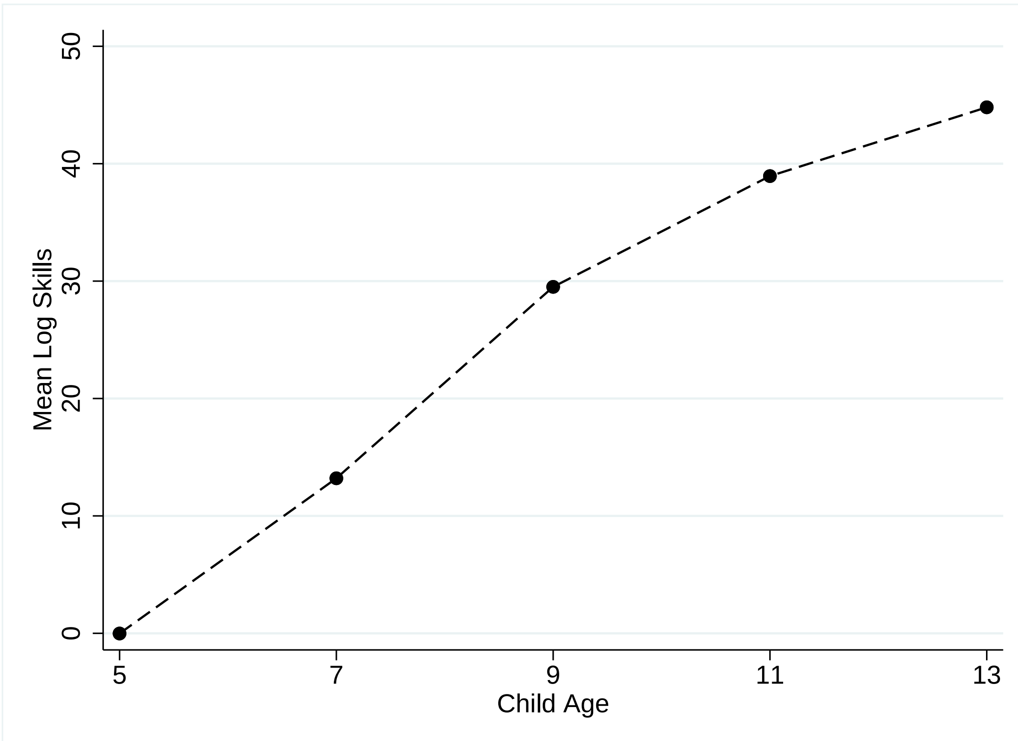
Notes: This figure shows the measurement error corrected estimates of the elasticity of children's skills at age 7-8 (θ_1) with respect to parental investments at age 5-6 (I_0): $\frac{\partial \ln \theta_1}{\partial \ln I_0} = \gamma_{2,0} + \gamma_{3,0} \ln \theta_0$.

Figure 2: Total Factor Productivity (TFP) Estimates



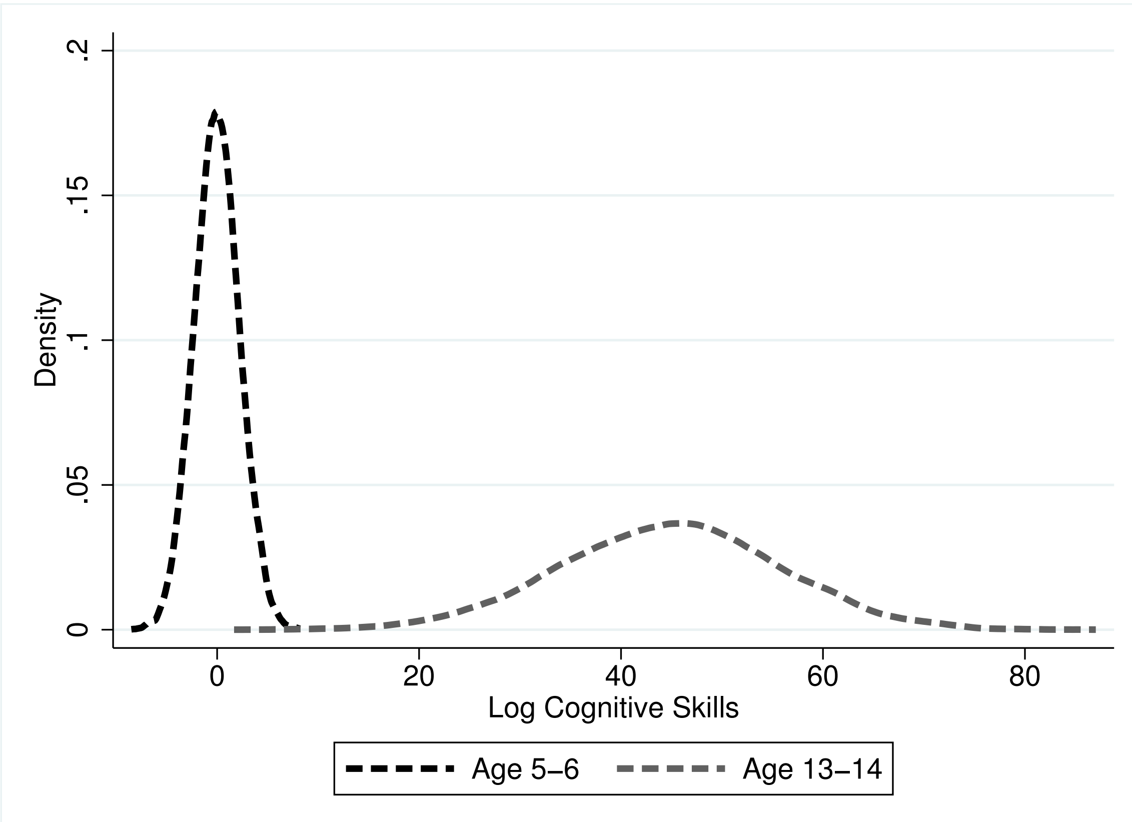
Notes: This figure shows the estimated log TFP (correcting for measurement error). The x-axis shows children age. Child age of 5 is age 5-6, 7 is age 7-8, and so on.

Figure 3: Estimated Mean of Log Latent Skills



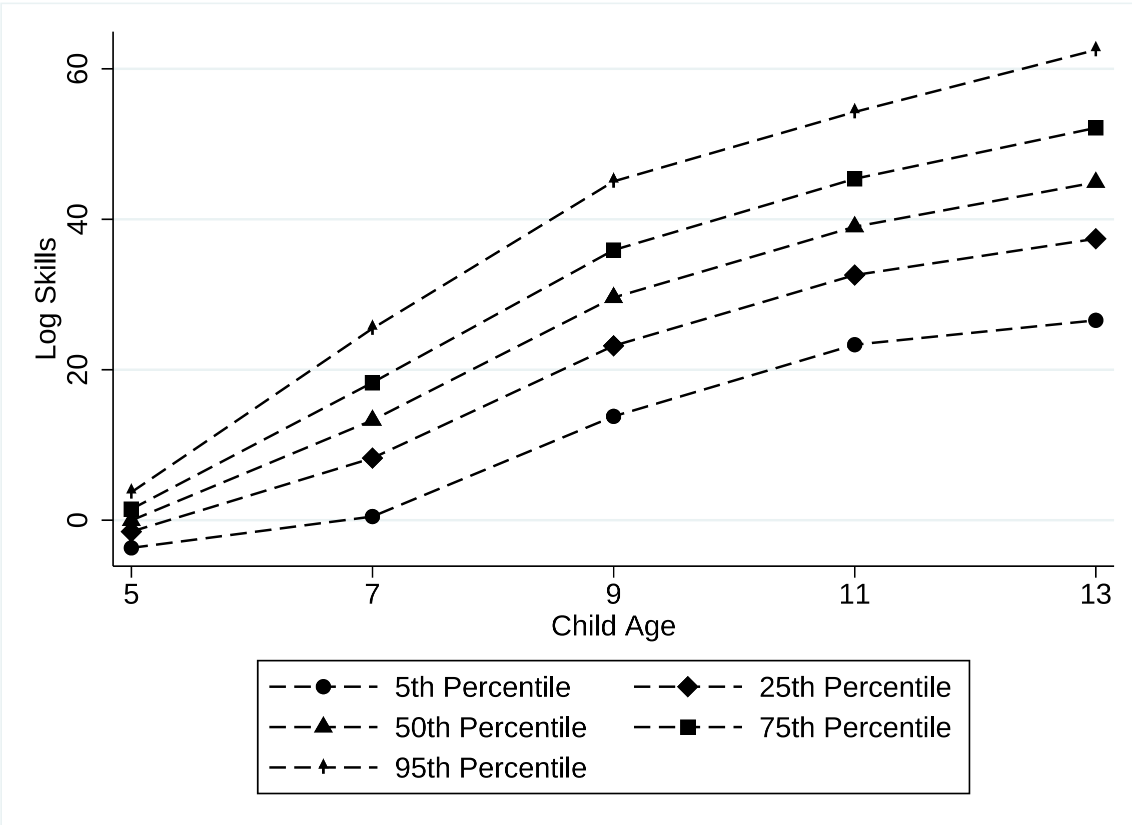
Notes: This figure provides the mean log latent skills ($E(\ln \theta_t)$) predicted by the estimated model, controlling for measurement error . The x-axis shows children age. Child age of 5 is age 5-6, 7 is age 7-8, and so on. Log latent skills at age 5-6 are normalized to be mean 0.

Figure 4: Estimated Distribution of Log Cognitive Latent Skills at Age 5-6 and Age 13-14



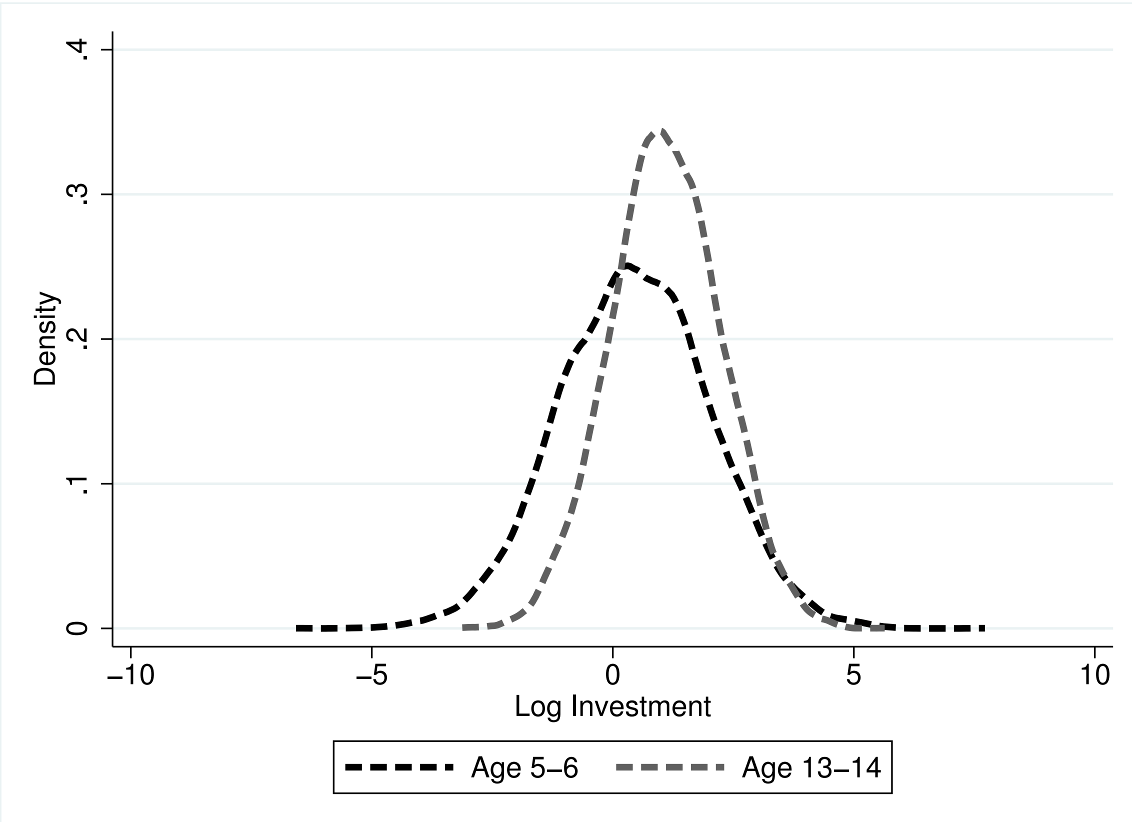
Notes: This figure shows the distribution of log latent skills at age 5-6 and at age 13-14 simulated from the estimated model, controlling for measurement error. Log latent skills at age 5-6 are normalized to be mean 0.

Figure 5: Estimated Dynamics in the Latent Skills Distribution



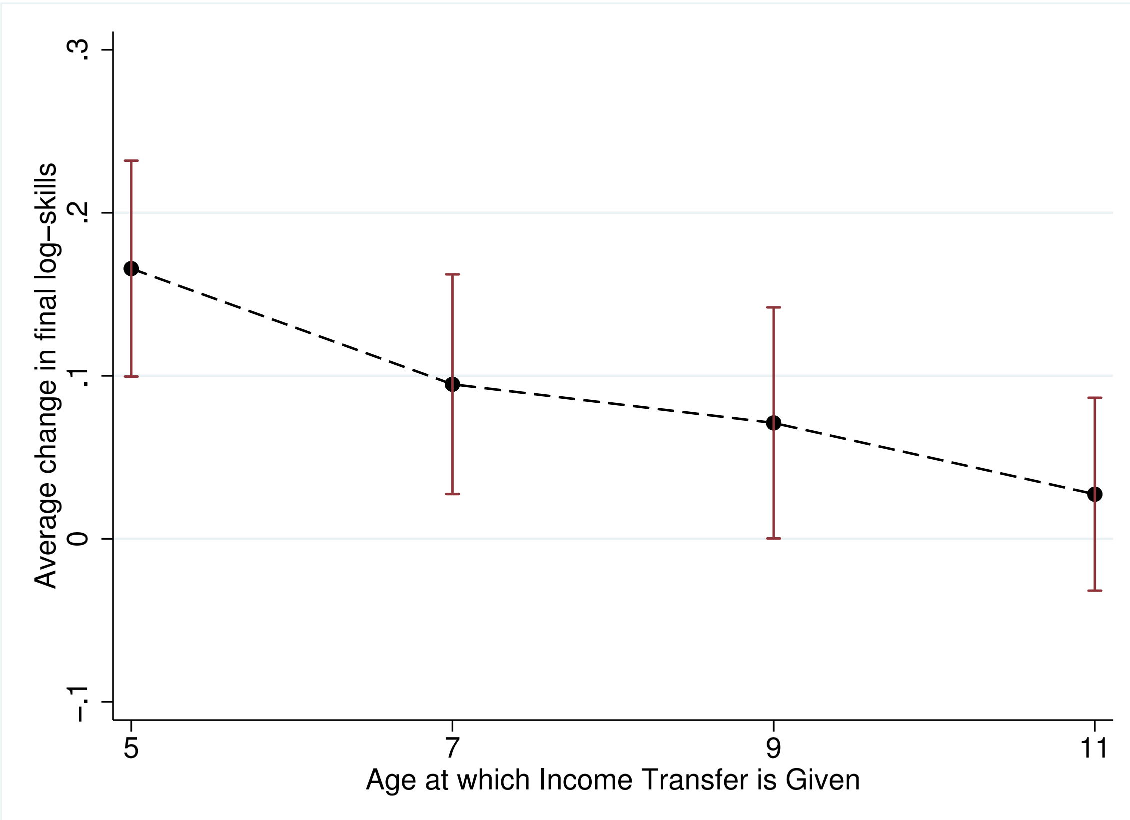
Notes: This figure shows the dynamics of the log latent skill distribution simulated from the estimated model, controlling for measurement error. Log latent skills at age 5-6 are normalized to be mean 0.

Figure 6: Estimated Distribution of Log Investments at Age 5-6 and Age 13-14



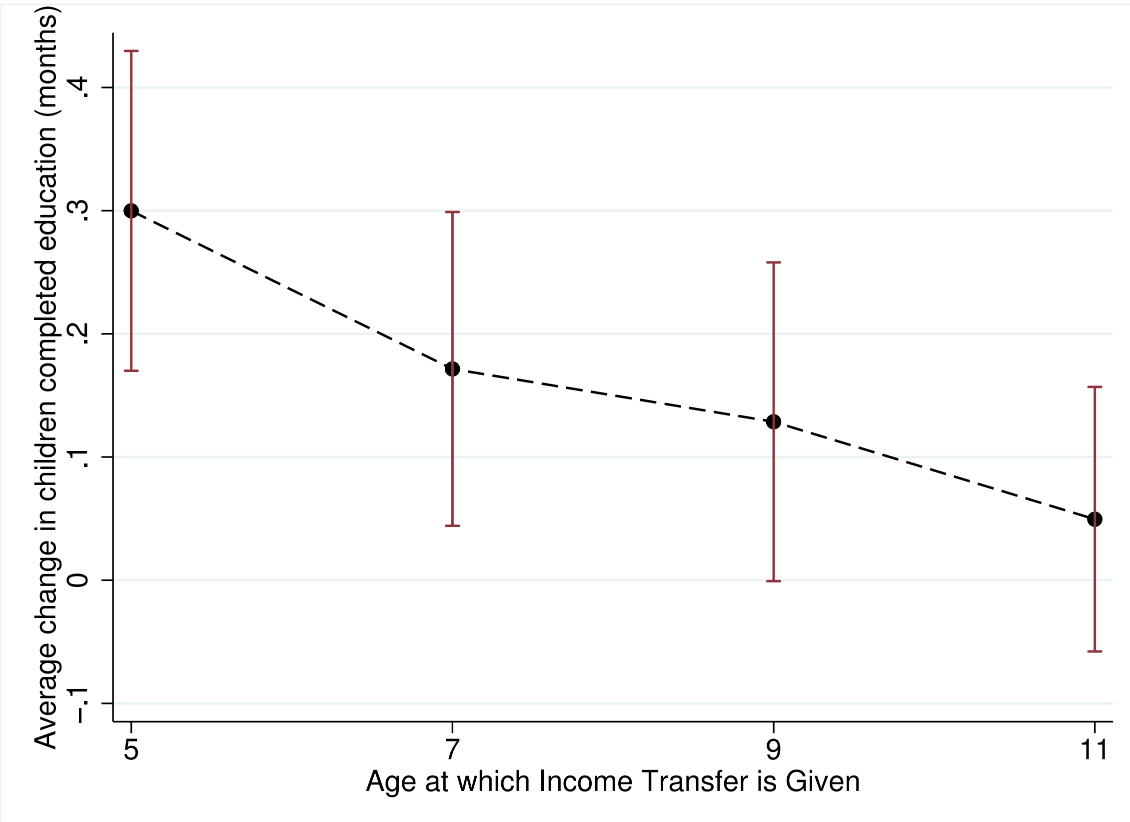
Notes: This figure shows the distribution of log latent investments at age 5-6 and at age 13-14 simulated from the estimated model, controlling for measurement error.

Figure 7: Average Effect of Income Transfer by Age of Transfer (Outcome: Final Period $\ln \theta_T$ Skills)



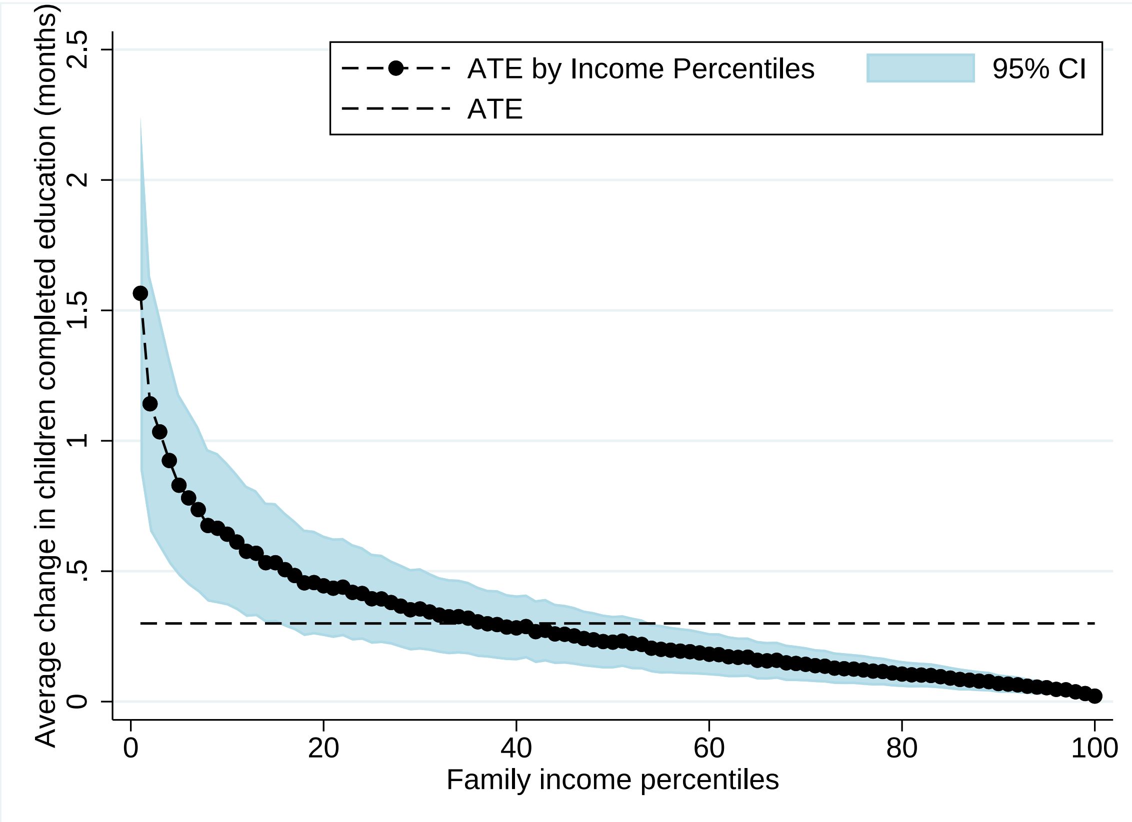
Notes: This figure shows the average log points change in the latent children's skills at age 13-14 by the different timing (age) of income transfer for the estimated model, controlling for measurement error. The transfer is \$10,000 in family income at some age t. We report $E\left[\ln \theta'_T(a) - \ln \theta_T\right]$, where $\theta'_T(a)$ is level of skill at age 13-14 with an income transfer of \$10,000 dollars provided to the family at age a, and θ_T is level of skill at age 13-14 at baseline (no income transfer).

Figure 8: Average Effect of an Income Transfer by Age of Transfer (Outcome: Schooling at Age 23)



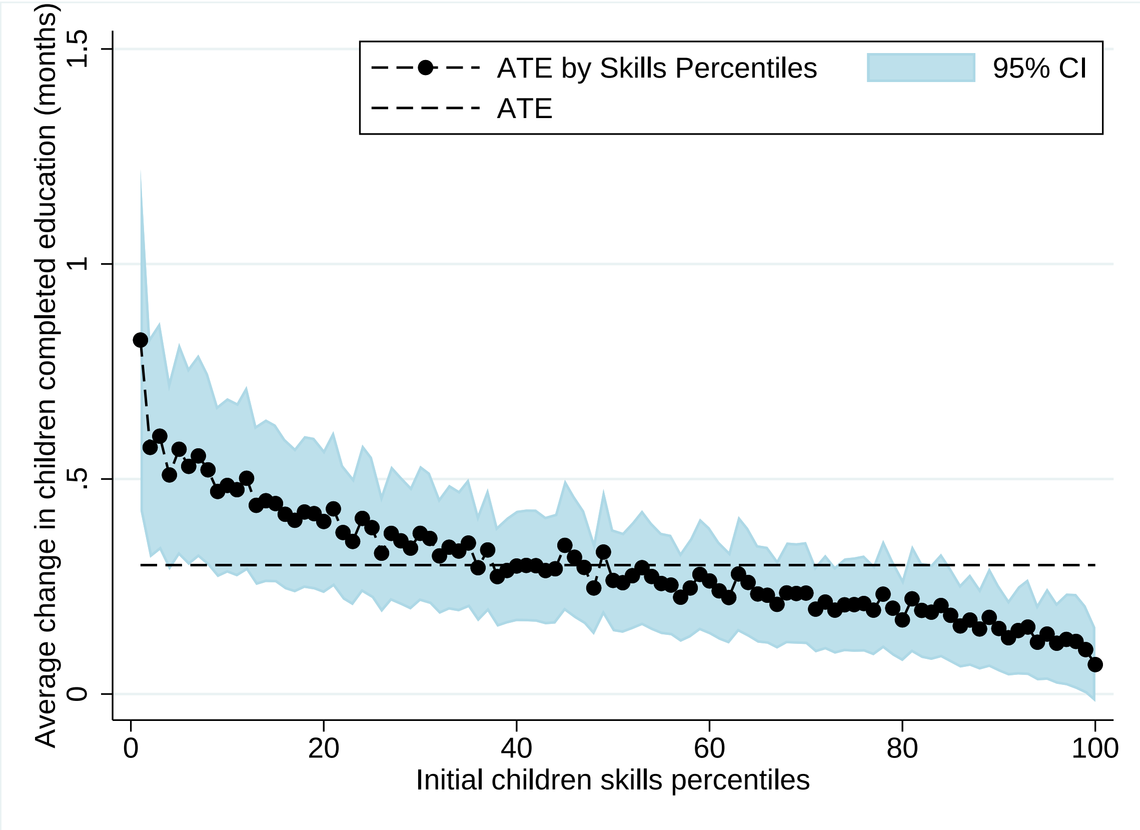
Notes: This figure shows the average change in the number of months of completed schooling at age 23 by different timing (age) of income transfer for the estimated model, controlling for measurement error. We report $E\left[S'(a) - S\right]$, where $S'(a)$ is the number of months of completed schooling at age 23 with an income transfer of \$10,000 given at age a while S is the number of months of completed schooling at baseline (no income transfer). This figure reports the results of the same policy experiment as Figure 7 but with respect to a different outcome measure (schooling).

Figure 9: Heterogeneity in Policy Effects by Age 5-6 Household Income (Outcome: Schooling at Age 23)



Notes: This figure plots the heterogeneous effect of a $10,000 income transfer at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income for the estimated model, controlling for measurement error. In the estimated income distribution for our sample, income categories 10, 50, and 90 contain families with about $14,000, $45,000, and $145,000 of annual family income. This figure also plots the average effect of the policy in the population

Figure 10: Heterogeneity in Policy Effects by Age 5-6 Children's Skills (Outcome: Schooling at Age 23)



Notes: This figure plots the heterogeneous effect of a $10,000 income transfer at age 5-6 on completed months of schooling by the percentile of the child's initial (age 5-6) skill for the estimated model, controlling for measurement error. This figure also plots the average effect of the policy in the population.

ONLINE APPENDIX

A Additional Analysis

A.1 Technologies and Output Elasticities

In this Appendix, we analyze the restrictions on key elasticities imposed by common functional forms. One rationale for the choice of a technology specification with free returns to scale is the flexibility this specification offers with respect to the implied output elasticity. We consider the output elasticity with respect to investment defined as

$$\epsilon_{I,t} \equiv \frac{\partial \ln \theta_{t+1}}{\partial \ln I_t}$$

This elasticity is key to quantifying the effects of policy interventions.

In the general CES case, with technology given by

$$\theta_{t+1} = \left[\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}},$$

the output elasticity is given by

$$\epsilon_{I,t} = \frac{\psi_t}{\phi_t}\left[\gamma \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}-1} \phi(1-\gamma_t) I_t^{\phi_t - 1} \cdot \frac{I_t}{\left[\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}}}$$

$$= \frac{\psi_t (1-\gamma_t) I_t^{\phi_t}}{\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}} \in [0, \infty)$$

In the special case of constant returns to scale (CRS), $\psi_t = 1$, and $\epsilon_{I,t} \in (0,1)$. CRS implies this elasticity is bounded from above by 1. The general free returns to scale case allows a larger than unit elastic response.

Similarly, the general translog function,

$$\ln \theta_{t+1} = \alpha_{1t} \ln \theta_t + \alpha_{2t} \ln I_t + \alpha_{3t} \ln \theta_t \ln I_t$$

with elasticity

$$\epsilon_{I,t} = \alpha_{1t} + \alpha_{3t} \ln \theta_t$$

also allows higher than unit elastic elasticities.

The CES technology with constant returns to scale restricts the output elasticity to be between 0 and 1: a one percent change in investment leads to a less than one percent change in next period skills. This prediction is independent of data, hence it can potentially be very restrictive in the context of child development and skill formation.

A.2 Semi-Parametric Identification

In this Appendix, we present a more general identification analysis, relaxing some of the functional form assumptions on the skill technology. Utilizing the main text log-linear specifications for measurement error, we begin with a measure of latent skill in period 1 $Z_{1,m}$ expressed as a function of both the measurement parameters and the technology:

$$\begin{aligned} Z_{\theta,1,m} &= \mu_{\theta,1,m} + \lambda_{\theta,1,m} \ln f_0(\theta_0, I_0) + (\lambda_{\theta,1,m}\eta_{\theta,0} + \epsilon_{\theta,1,m}) \\ &= q_0(\theta_0, I_0) + u_{0,m} \end{aligned} \tag{A-1}$$

where the combined residual $u_{0,m} = \lambda_{\theta,1,m}\eta_{\theta,0} + \epsilon_{\theta,1,m}$ is mean-zero. The period 0 child's skill (θ_0) and parental investments (I_0) on the RHS are unobserved, but we have some "error-contaminated" measurements $(\widetilde{Z}_{\theta,0,m}, \widetilde{Z}_{I,0,m})$ derived from the previous identification step: $\widetilde{Z}_{\theta,0,m} = \ln\theta_0 + \widetilde{\epsilon}_{\theta,0,m}$ and $\widetilde{Z}_{I,0,m} = \ln I_0 + \widetilde{\epsilon}_{I,0,m}$, where $\widetilde{\epsilon}_{\theta,0,m} = \frac{\epsilon_{\theta,0,m}}{\lambda_{\theta,0,m}}$ and $\widetilde{\epsilon}_{I,0,m} = \frac{\epsilon_{I,0,m}}{\lambda_{I,0,m}}$.

Equation (A-1) can be thought of as a semi-parametric regression equation relating an observed measure of period 1 skills to a non-parametric function of unobserved period 0 skills and investment and an additively separable error term. The new error term in this equation $u_{0,m}$ has two parts: the production shock $\eta_{\theta,0}$ and the measurement error $\epsilon_{\theta,1,m}$. There are two identification challenges here: (i) the unobservability of the RHS skills and investents (θ_0, I_0); and (ii) the potential endogeneity of these inputs: the error term $u_{0,m}$ being correlated with $\ln\theta_0$ and $\ln I_0$. As noted in Adusumilli and Otsu (2018), estimating the model given in (A-1) relies on two longstanding and largely parallel econometric research programs on non-parametric IV models and models with errors-in-variables (e.g., Hausman et al., 1991; Schennach, 2004). Sufficient conditions for identification of q_0 are given in Adusumilli and Otsu (2018), and they rely on the existence of a relevant instrumental variable vector W_0 which satisfy two key conditions: (i) $E(u_{0,m}|W_0) = 0$ and (ii) $(\widetilde{\epsilon}_{\theta,0,m}, \widetilde{\epsilon}_{I,0,m}) \perp W_0$. In addition, a third requirement on the measurement errors is (iii) $(\widetilde{\epsilon}_{\theta,0,m}, \widetilde{\epsilon}_{I,0,m}) \perp (\theta_0, I_0)$.

Given the non-parametric identification of the function q_0, the next step is unpacking the components of this function to provide identification for the dynamics of the next period skills. Without loss of generality, we write the first term of the production technology in (2) as

$$\ln f_0(\theta_0, I_0) = \ln A_0 + \psi_0 \ln K_0(\theta_0, I_0), \tag{A-2}$$

where $\ln A_0$ and ψ_0 are the location and scale of the (log) technology, and $K_0(\theta_0, I_0)$ is a Known Location and Scale (KLS) function, which we define as follows:

Definition A-1 *A production function $K_t(\theta_t, I_t)$ has* Known Location and Scale (KLS) *if for two non-zero input vectors (θ_t', I_t') and (θ_t'', I_t''), where the input vectors are distinct, the outputs $K_t(\theta_t', I_t')$ and $K_t(\theta_t'', I_t'')$ (with $K_t(\theta_t', I_t') \neq K_t(\theta_t'', I_t'')$) are both known (do not depend on unknown parameters), finite, and non-zero.*

A production technology with known location and scale implies that, for a change in inputs from (θ_t', I_t') to (θ_t'', I_t''), the change in output $K_t(\theta_t', I_t') - K_t(\theta_t'', I_t'')$ is known. Other points in the production possibilities set may be unknown, i.e. they depend on free parameters to be estimated. Writing the technology as in (A-2), we have intuitively separated out two parameters representing location and scale from the general function f_0, the parameters $\ln A_0$ and ψ_0. A leading example of a KLS function is the CRS CES function:

$$\theta_{t+1} = (\gamma_t \theta_t^{\sigma_t} + (1-\gamma_t) I_t^{\sigma_t})^{1/\sigma_t}$$

with $\gamma_t \in (0,1)$ and $\sigma_t \in (-\infty, 1]$. In this case it is easy to show that, for all pairs (θ_t, I_t) such that $\theta_t = I_t$, the output is known: $\theta_{t+1} = \theta_t = I_t$. This result follows from the constant return to scale property of the CES. Suppose $\theta_t = I_t = a$, we then have:

$$\theta_{t+1} = (\gamma_t a^{\sigma_t} + (1-\gamma_t) a^{\sigma_t})^{1/\sigma_t} = (a^{\sigma_t})^{1/\sigma_t} = a\,.$$

Returning to the general problem, substituting equation (A-2) into the main equation (A-1), we have

$$Z_{1,m} = (\mu_{\theta,1,m} + \ln A_0 \lambda_{\theta,1,m}) + (\lambda_{\theta,1,m}\psi_0) \ln K_0(\theta_0, I_0) + u_{0,m} \tag{A-3}$$

At this point, we cannot separately identify the period 1 measurement parameters $(\mu_{\theta,1,m}, \lambda_{\theta,1,m})$ from production function parameters $(\ln A_0, \psi_0)$. That is, we cannot separately identify the location and scale of the measurement function from the location and scale of the production function. We consider identification under one of two prototypical restrictions:

Assumption A-1 *Measurement Function Restriction*

$\mu_{\theta,t,m} = \mu_{\theta,0,m}$ *and* $\lambda_{\theta,t,m} = \lambda_{\theta,0,m}$ *for all* $t > 0$ *and for some* m

Assumption A-2 *Production Function Restriction*

$\ln A_t = 0$ *and* $\psi_t = 1$ *for all* t

Under *either* set of restrictions, we identify all of the parameters of interest. Let (θ_0', I_0') and (θ_0'', I_0'') be the two input vectors for which K_0 is, by definition, known and non-zero. Then we have a system of two equations:

$$q_0(\theta_0', I_0') = (\mu_{\theta,1,m} + \ln A_0 \lambda_{\theta,1,m}) + (\lambda_{\theta,1,m}\psi_0) \ln K_0(\theta_0', I_0')$$

$$q_0(\theta_0'', I_0'') = (\mu_{\theta,1,m} + \ln A_0 \lambda_{\theta,1,m}) + (\lambda_{\theta,1,m}\psi_0) \ln K_0(\theta_0'', I_0'')$$

where the q_0 function (defined in equation A-1) is identified given the arguments above. The two parameters $\alpha' = (\mu_{\theta,1,m} + \ln A_0 \lambda_{\theta,1,m})$ and $\alpha'' = (\lambda_{\theta,1,m}\psi_0)$ are just-identified. Under the measurement function restriction Assumption A-1, we identify $\ln A_0$ and ψ_0, in addition to the other parameters. Or, under the production function restriction Assumption A-2, we identify $\mu_{\theta,1,m}$ and $\lambda_{\theta,1,m}$, in addition to the other parameters.

Importantly, one does not need to assume *both* a production function and measurement function restriction. Either assumption is sufficient for identification in this context, and imposing both is over-identifying. There are other types of restrictions that would be sufficient for identification, but these two broad classes of restrictions help clarify an important range of options.

A.3 A Reappraisal of the Value-added Analysis

A large and growing literature estimates teacher quality by measuring teachers' "productivity" in affecting student test scores (Chetty et al., 2014). Rare in this literature is data allowing teacher quality to be anchored to adult outcomes (for an exception see Chetty et al., 2011), and almost the entirety of the empirical results rest on particular measures of student skills, typically grade-specific standardized test scores (for some discussion of general issues see Ballou, 2009).

Although it is not common in the current literature to express teacher value-added in terms of latent child skills, we can write the standard framework in our notation as

$$\ln \theta_{i,t+1} = \ln A_t + \gamma_t \ln \theta_{i,t} + \delta_{j(i),t} + \eta_{i,\theta,t} \tag{A-4}$$

where $j(i)$ indicates that teacher j teaches student i, t is grade-level rather than age, and $\delta_{j(i),t}$ represents teacher j's "value-added". Here $\delta_{j(i),t}$ replaces parental investment in our models, where the productivity of teachers $\delta_{j(i),t}$ is treated as an unobserved fixed effect common to a classroom of students. For recent work estimating latent factor models including both parental and school inputs, see Agostinelli et al. (2019).

The current value-added literature typically estimates models of this form, using observed measures for student test scores:

$$Z_{i,\theta,t+1,m} = \beta_{0,t,m} + \beta_{1,t,m} Z_{i,\theta,t,m} + \widetilde{\delta}_{j(i),t,m} + \pi_{i,\theta,t,m} \tag{A-5}$$

where $Z_{i,\theta,t+1,m}$ and $Z_{i,\theta,t,m}$ are measures of the underlying latent skills (typically age-standardized test scores in mathematics or reading). $\widetilde{\delta}_{j(i),t,m}$ is the value-added of teacher j on grade $t+1$, given the particular left-hand-side measure of skills m in $t+1$. By using the measurement model in (4), we can map the parameters in (A-5) into the parameters of interest in (A-4):

$$\beta_{1,t,m} = \gamma_{t,m} \cdot \frac{\lambda_{\theta,t+1,m}}{\lambda_{\theta,t,m}}$$
$$\widetilde{\delta}_{j(i),t,m} = \delta_{j(i),t} \cdot \lambda_{\theta,t+1,m}$$

The above result highlights two main conclusions. First, in the absence of additional restrictions on the measurement model–for example our age-invariance restrictions–the estimated educational production function parameter $\beta_{1,t,m}$ does not have a well-defined interpretation, as the estimated value depends on the changes in how latent skills influence the observed measures (factor loadings) between the grade levels t and $t+1$.

Second, the main statistic of interest–the variance of the the estimated teacher effects, which indicates how "productive" teachers are in affecting student learning–is not a scale-free parameter:

$$V(\widetilde{\delta}_{j(i),t,m}) = V(\delta_{j(i),t}) \cdot \lambda^2_{\theta,t+1,m}$$

Higher or lower estimates of the variance in teacher effects can be due to the particular factor loadings of the observed scores m, distinct from the variance of the actual underlying teacher value-added. Although the common scaling parameter does not affect the *rankings* of teachers within grade, if this scale changes across-grades, then one cannot separately identify across-grade changes in the true productivity of teachers from changes in the scale of the measures.

B Additional Tables and Figures

B.1 Additional Tables for Estimates Corrected for Measurement Error

Table B-1: Estimates for Initial Conditions

	Log Child Skills at age 5	Log Mother Cognitive Skills	Log Mother Noncognitive Skills	Log Family Income
	Variance-Covariance Matrix			
Log Child Skills at age 5	4.947 (0.469)	6.254 (0.476)	0.122 (0.031)	0.668 (0.065)
Log Mother Cognitive Skills	6.254 (0.476)	30.190 (1.027)	0.593 (0.136)	2.588 (0.098)
Log Mother Noncognitive Skills	0.122 (0.031)	0.593 (0.136)	0.046 (0.017)	0.058 (0.012)
Log Family Income	0.668 (0.065)	2.588 (0.098)	0.058 (0.012)	0.780 (0.018)
	Correlation Matrix			
Log Child Skills at age 5	1.000 (0.000)	0.512 (0.026)	0.256 (0.029)	0.340 (0.027)
Log Mother Cognitive Skills	0.512 (0.026)	1.000 (0.000)	0.504 (0.025)	0.533 (0.015)
Log Mother Noncognitive Skills	0.256 (0.029)	0.504 (0.025)	1.000 (0.000)	0.307 (0.022)
Log Family Income	0.340 (0.027)	0.533 (0.015)	0.307 (0.022)	1.000 (0.000)

Notes: This table shows the estimated variance-covariance matrix (Σ_Ω) and associate correlation matrix of the initial conditions at age 5-6. Standard errors in parenthesis are computed using a clustered bootstrap at the family level.

Table B-2: Estimates for Investment Equation

Model	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.230	0.027	0.020	0.018
at age 5	(0.059)	(0.009)	(0.009)	(0.009)
	[0.15,0.33]	[0.02,0.04]	[0.01,0.04]	[0.01,0.04]
Log Mother	0.071	0.004	0.012	-0.005
Cognitive Skills	(0.022)	(0.009)	(0.015)	(0.013)
	[0.04,0.12]	[-0.01,0.02]	[-0.01,0.04]	[-0.03,0.02]
Log Mother	0.359	0.742	0.694	0.712
Noncognitive Skills	(0.130)	(0.059)	(0.083)	(0.087)
	[0.11,0.53]	[0.64,0.82]	[0.53,0.80]	[0.54,0.81]
Log Family	0.341	0.227	0.274	0.275
Income	(0.076)	(0.056)	(0.076)	(0.087)
	[0.25,0.48]	[0.16,0.32]	[0.17,0.42]	[0.17,0.45]
Standard Deviation	1.186	1.019	0.868	1.087
Shocks	(0.230)	(0.147)	(0.235)	(0.295)
	[0.97,1.54]	[0.83,1.29]	[0.66,1.32]	[0.82,1.72]

Notes: This table shows the measurement error corrected estimates for the investment equation. Each column shows the coefficients of the investment equation at the given ages. The dependent variable is (log) investment in period t, determined by the RHS variables at time t. For example, the first column shows the coefficients at age 5-6 parental investments determined by age 5-6 child's skill and family income. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

Table B-3: Estimates for Adult Outcome Equation

Model	Schooling	Log Wage
Constant	7.088	8.394
	(0.397)	(0.252)
	[6.56,7.71]	[8.03,8.79]
Log Children Skills at age 13-14	0.151	0.041
	(0.009)	(0.006)
	[0.14,0.16]	[0.03,0.05]
Variance Shocks	4.333	0.875
	(0.142)	(0.064)
	[4.07,4.56]	[0.77,0.97]

Notes: This table shows the estimates for two adult outcome equation specifications: schooling and log earnings. In both cases the estimates are corrected for measurement error. The dependent variable is either the years of completed education for the child at age 23 or log earnings at age 29. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

Table B-4: Estimates for Income Process

Constant	0.377
	(0.012)
	[0.36,0.40]
Log Family Income t-1	0.753
	(0.008)
	[0.74,0.76]
Standard Deviation Innovation	0.579
	(0.008)
	[0.57,0.59]
Initial Mean Log-Income	1.372
at age 5-6	(0.016)
	[1.35,1.40]

Notes: This table shows the estimates for the income process. The dependent variable is log family income at time t. Log Family Income $t-1$ is log family income two years prior. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

Table B-5: Children's Skills Measures

Measures	Range Values	Age Range	Scoring Order
(The Peabody Individual Achievement Test):			
Math	0-84	5-14	Positive
Recognition	0-84	5-14	Positive
Comprehensive	0-84	5-14	Positive

Notes: This table shows the features of children cognitive measures. The first column indicate each type of children skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The third column shows the minimum and maximum children age at which each measure is available. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B-6: Mothers Cognitive Skills Measures

Measures	Range Values	Scoring Order
Arithmetics	0-30	Positive
Word Knowledge	0-35	Positive
Paragraph Composition	0-15	Positive
Numeric Operations	0-50	Positive
Coding Speed	0-84	Positive
Math Knowledge	0-25	Positive

Notes: This table shows the features of mother cognitive measures. The first column indicate each type of mother cognitive skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B-7: Mothers Noncognitive Skills Measures

Type of variables	Range Values	Label	Scoring Order
Mother Noncognitive Measures			
(Rosenberg indexes): I am a person of worth I have a number of good qualities I am able to do things as well as most other people I take a positive attitude toward myself	1-4	1= Strongly agree 2= Agree 3= Disagree 4= Strongly disagree	Negative
I am inclined to feel that I am a failure I felt I do not have much to be proud of I wish I could have more respect for myself I certainly feel useless at times At times I think I am no good at all	1-4	1= Strongly agree 2= Agree 3= Disagree 4= Strongly disagree	Positive
(Rotter Indexes):			
Rotter 1 (Life is in control or not)	1-4	1= In Control and closer to my opinion 2= In control but slightly closer to my opinion 3= Not in control but slightly closer to my opinion 4= Not in control and closer to my opinion	Negative
Rotter 2 (Plans work vs Matter of Luck)	1-4	1= Plans work and closer to my opinion 2= Plans work but slightly closer to my opinion 3= Matter of Luck but slightly closer to my opinion 4= Matter of Luck and closer to my opinion	Negative
Rotter 3 (Luck not a factor vs Flip a coin)	1-4	1= Luck not a factor and closer to my opinion 2= Luck not a factor but slightly closer to my opinion 3= Flip a coin but slightly closer to my opinion 4= Flip a coin and closer to my opinion	Negative
Rotter 4 (Luck big role vs Luck no role)	1-4	1= Luck big role and closer to my opinion 2= Luck big role but slightly closer to my opinion 3= Luck no role but slightly closer to my opinion 4= Luck no role and closer to my opinion	Positive

Notes: This table shows the features of mother noncognitive measures. The first column indicate each type of mother cognitive skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The third column shows the type of answers associated with each measure value. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B-8: Descriptive Statistics about Children's Cognitive Skills Measures

Measures	Mean	Std	Min	Max	Number of Values
Age 5-6					
PIAT Math	11.858	4.278	0.000	37.000	32.000
PIAT Recognition	12.864	5.048	0.000	57.000	35.000
PIAT Comprehensive	12.770	4.930	0.000	49.000	35.000
Age 7-8					
PIAT Math	23.016	8.681	0.000	74.000	58.000
PIAT Recognition	25.748	8.774	0.000	80.000	67.000
PIAT Comprehensive	24.099	8.142	0.000	69.000	60.000
Age 9-10					
PIAT Math	38.720	10.832	0.000	84.000	71.000
PIAT Recognition	40.825	11.487	0.000	84.000	76.000
PIAT Comprehensive	37.540	10.231	0.000	78.000	64.000
Age 11-12					
PIAT Math	48.184	10.543	0.000	84.000	78.000
PIAT Recognition	51.079	13.278	0.000	84.000	74.000
PIAT Comprehensive	45.732	11.272	0.000	84.000	72.000
Age 13-14					
PIAT Math	53.767	11.387	0.000	84.000	78.000
PIAT Recognition	58.670	14.262	0.000	84.000	74.000
PIAT Comprehensive	51.015	12.229	0.000	84.000	74.000

Notes: This table shows main sample statistics of children cognitive skills measures by children age.

Table B-9: Descriptive Statistics of Parental Investment Measures

	Parental Investments			
Measures	Mean	Std	Min	Max
How often mom reads to child	4.22	1.41	1	6
How often mom eats with child	3.32	1.61	0	5
How often child was taken to museum	2.19	0.97	1	5
How often child is praised	5.56	4.37	0	20
How often complimented child	4.70	4.05	0	20

Notes: This table shows main sample statistics of parental investment measures.

Table B-10: Descriptive Statistics of Mother Cognitive and Noncognitive Skills Measures

Mother Cognitive Skills					
Measures	Mean	Std	Min	Max	Number of Values
Mom‘s Arithmetic Reasoning Test Score	13.946	6.603	0.000	30.000	31.000
Mom‘s Word Knowledge Test Score	21.773	8.562	0.000	35.000	36.000
Mom‘s Paragraph Composition Test Score	9.620	3.778	0.000	15.000	16.000
Mom‘s Numerical Operations Test Score	31.044	11.831	0.000	50.000	51.000
Mom‘s Coding Speed Test Score	42.953	17.468	0.000	84.000	85.000
Mom‘s Mathematical Knowledge Test Score	10.853	5.867	0.000	25.000	26.000
Mother Non Cognitive Skills					
Mom‘s Self-Esteem: ”I am a person of worth”	2.461	0.549	0.000	3.000	4.000
Mom‘s Self-Esteem: ” I have good qualities”	2.338	0.539	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I am a failure”	3.379	0.618	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I am as capable as others”	2.291	0.567	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I have nothing to be proud of”	3.360	0.669	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I have a positive attitude”	2.183	0.619	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I wish I had more self-respect”	2.796	0.817	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I feel useless at times”	2.650	0.770	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I sometimes think I am no good”	3.039	0.802	1.000	4.000	4.000
Mom‘s Rotter Score:”I have no control”	2.863	1.058	1.000	4.000	4.000
Mom‘s Rotter Score: ”I make no plans for the future”	2.386	1.192	1.000	4.000	4.000
Mom‘s Rotter Score: ”Luck is big factor in life”	3.205	0.856	1.000	4.000	4.000
Mom‘s Rotter Score: ”Luck plays big role in my life”	2.594	1.024	1.000	4.000	4.000

Notes: This table shows main sample statistics of mother cognitive and non-cognitive skills measures.

Table B-11: Measurement Parameter Estimates for Children's Cognitive Measures

Measures	μ	λ	Signal	Noise
Age 5-6				
PIAT Math	11.858	1.000	0.273	0.727
PIAT Recognition	12.864	2.238	0.981	0.019
PIAT Comprehensive	12.770	2.159	0.957	0.043
Age 7-8				
PIAT Math	11.858	1.000	0.764	0.236
PIAT Recognition	15.592	0.906	0.613	0.387
PIAT Comprehensive	15.014	0.802	0.559	0.441
Age 9-10				
PIAT Math	11.858	1.000	0.777	0.223
PIAT Recognition	10.297	1.136	0.892	0.108
PIAT Comprehensive	12.273	0.936	0.763	0.237
Age 11-12				
PIAT Math	11.858	1.000	0.804	0.196
PIAT Recognition	2.107	1.347	0.920	0.080
PIAT Comprehensive	6.129	1.089	0.834	0.166
Age 13-14				
PIAT Math	11.858	1.000	0.924	0.076
PIAT Recognition	8.556	1.195	0.842	0.158
PIAT Comprehensive	9.041	1.002	0.803	0.197

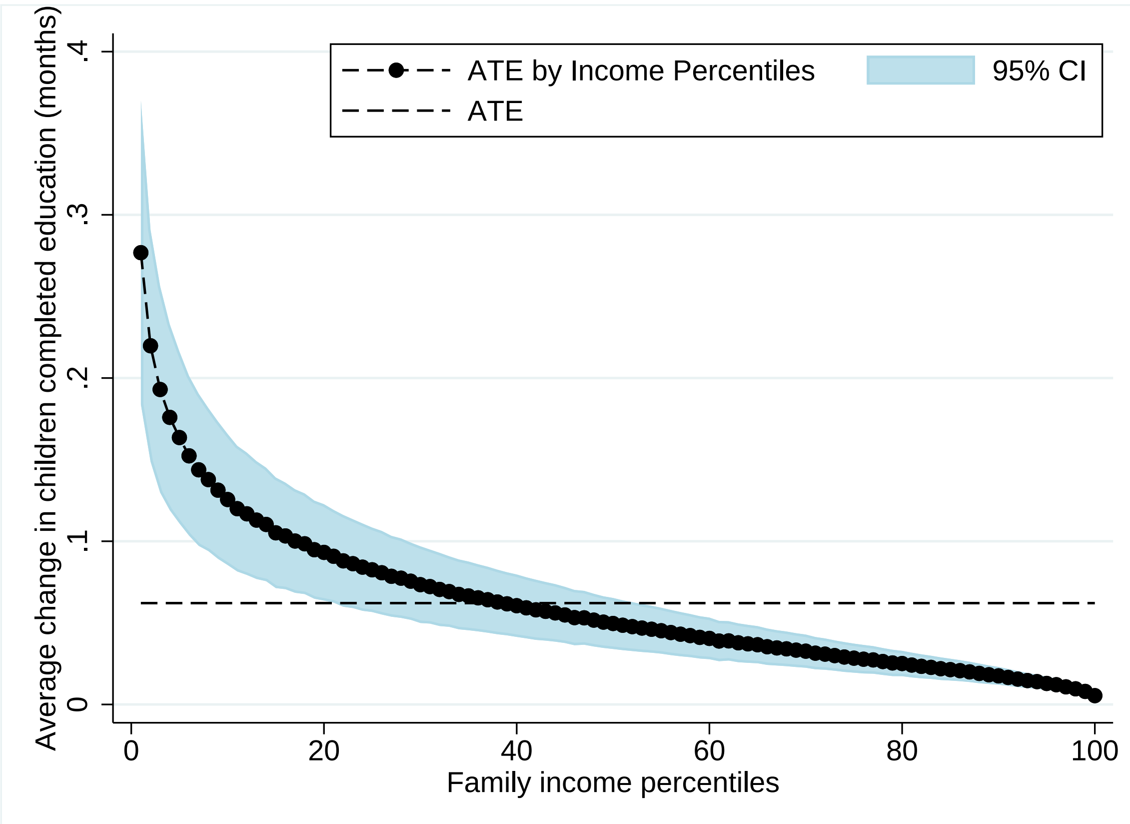
Notes: This table shows the measurement error parameters and associated statistics for children cognitive measures. The first two columns shows the measurement parameters (μ and λ), while the last two columns shows the signal and noise variance decomposition for the children cognitive measures.

Table B-12: Measurement Parameter Estimates for Mother Cognitive and Noncognitive Measures

Measures	μ	λ	Signal	Noise
Mother Cognitive Skills				
Mom's Arithmetic Reasoning	13.946	1.000	0.692	0.308
Mom's Word Knowledge	21.773	1.345	0.745	0.255
Mom's Paragraph Composition	9.620	0.584	0.722	0.278
Mom's Operations	31.044	1.720	0.638	0.362
Mom's Coding Speed	42.953	2.308	0.527	0.473
Mom's Mathematical Knowledge	10.853	0.854	0.639	0.361
Mother Non Cognitive Skills				
Mom's Self-Esteem: "I am a person of worth"	2.461	1.000	0.152	0.848
Mom's Self-Esteem: "I have good qualities"	2.338	1.263	0.252	0.748
Mom's Self-Esteem: "I am a failure"	3.379	1.612	0.311	0.689
Mom's Self-Esteem: "I am as capable as others"	2.291	1.127	0.181	0.819
Mom's Self-Esteem: "I have nothing to be proud of"	3.360	1.746	0.312	0.688
Mom's Self-Esteem: "I have a positive attitude"	2.183	1.474	0.260	0.740
Mom's Self-Esteem: "I wish I had more self-respect"	2.796	2.080	0.297	0.703
Mom's Self-Esteem: "I feel useless at times"	2.650	1.861	0.268	0.732
Mom's Self-Esteem: "I sometimes think I am no good"	3.039	2.096	0.313	0.687
Mom's Rotter Score: "I have no control"	2.863	1.497	0.092	0.908
Mom's Rotter Score: "I make no plans for the future"	2.386	2.081	0.140	0.860
Mom's Rotter Score: "Luck is a big factor in life"	3.205	1.372	0.118	0.882
Mom's Rotter Score: "Luck plays big role in my life"	2.594	1.002	0.044	0.956

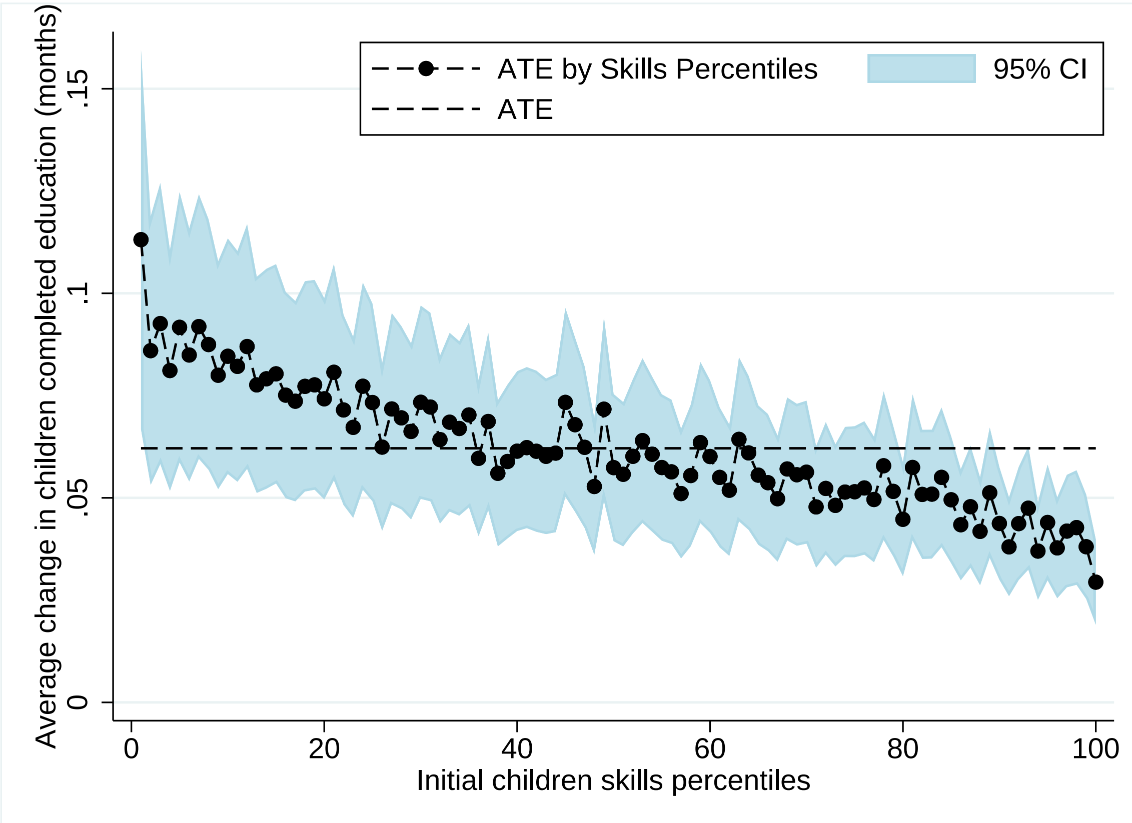
Notes: This table shows the measurement error parameters and associated statistics for mother cognitive and noncognitive measures. The first two columns show the measurement parameters (μ and λ), while the last two columns show the signal and noise variance decomposition for the mother measures.

Figure B-1: Heterogeneity in Policy Effects by Age 5 Household Income (Outcome: Schooling at Age 23, No Measurement Error Correction)



Notes: This figure plots the heterogeneous effect of a $10,000 income transfer at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income for the estimated model, not controlling for measurement error. In the estimated income distribution for our sample, income categories 10, 50, and 90 contain families with about $14,000, $45,000, and $145,000 of annual family income. This figure also plots the average effect of the policy in the population.

Figure B-2: Heterogeneity in Policy Effects by Age 5 Children's Skills (Outcome: Schooling at Age 23, No Measurement Error Correction)



Notes: This figure plots the heterogeneous effect of a $10,000 income transfer at age 5-6 on completed months of schooling by the percentile of the child's initial (age 5-6) skill for the estimated model, not controlling for measurement error. This figure also plots the average effect of the policy in the population.

C Additional Robustness Analysis

Table C-1: Estimates for Skill Technology with CHS Sample

Model	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	1.638	0.879	0.866	0.988
	(0.261)	(0.056)	(0.072)	(0.056)
	[1.17,1.99]	[0.76,0.94]	[0.74,0.98]	[0.88,1.07]
Log Investments	1.120	-0.141	-0.476	-0.602
	(0.602)	(0.283)	(0.646)	(0.591)
	[0.42,2.21]	[-0.64,0.29]	[-1.29,0.26]	[-1.42,0.51]
Log Skills ×	-0.057	0.014	0.025	0.012
Log Investments	(0.153)	(0.015)	(0.022)	(0.014)
	[-0.26,0.22]	[-0.00,0.04]	[-0.00,0.06]	[-0.01,0.03]
Standard Deviation	5.459	3.684	3.536	3.624
Shocks	(0.301)	(0.331)	(0.358)	(0.364)
	[5.06,6.03]	[3.27,4.46]	[3.11,4.29]	[3.32,4.44]
Log TFP	14.057	17.928	12.825	7.214
	(0.661)	(0.723)	(1.907)	(1.882)
	[12.88,14.94]	[17.16,19.42]	[10.22,16.99]	[3.98,11.24]

Notes: This table shows the measurement error corrected estimates for the technology of skills formation for the same estimating sample as in Cunha et al. (2010): firstborn white children. Each column shows the coefficients of the technology of skills formation at the given age. The dependent variable is log skills in the next period $t+1$, and the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

Table C-2: Estimates for Skill Technology with Additional Controls

Model	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	1.920	1.084	0.896	1.067
	(0.138)	(0.036)	(0.025)	(0.027)
	[1.68,2.15]	[1.02,1.15]	[0.85,0.93]	[1.02,1.11]
Log Investments	0.745	0.673	0.713	0.270
	(0.250)	(0.333)	(0.390)	(0.557)
	[0.41,1.11]	[0.15,1.28]	[-0.03,1.18]	[-0.53,1.19]
Log Skills ×	-0.098	-0.004	-0.003	0.003
Log Investments	(0.062)	(0.018)	(0.013)	(0.010)
	[-0.21,-0.03]	[-0.04,0.03]	[-0.02,0.02]	[-0.02,0.02]
Standard Deviation	5.612	4.519	3.585	4.019
Shocks	(0.173)	(0.184)	(0.180)	(0.246)
	[5.38,5.93]	[4.28,4.88]	[3.27,3.87]	[3.71,4.43]
Log TFP	13.420	15.060	12.105	3.133
	(0.303)	(0.431)	(0.567)	(0.942)
	[12.95,13.96]	[14.36,15.82]	[11.35,13.18]	[1.52,4.74]

Notes: This table shows the measurement error corrected estimates for the technology of skills formation once we add additional controls in the investment equation in (8). In particular, we control for: a set of dummies for the maximum number of children ever observed in the household, a set of dummies for the mother's marital status, maternal hours worked, maternal hourly wage and a dummy for employment status (employed/non-employed). Each column shows the coefficients of the technology of skills formation at the given age. The dependent variable is log skills in the next period $t + 1$, and the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8. Both standard errors in parenthesis and the 90% confidence interval in square brackets are computed using a clustered bootstrap at the family level.

D Skills measures in CNLSY79

Measures for Cognitive Skills

- **Peabody Picture Vocabulary Test**

  The Peabody Picture Vocabulary Test, revised edition (PPVT-R) "measures an individual's receptive (hearing) vocabulary for Standard American English and provides, at the same time, a quick estimate of verbal ability or scholastic aptitude" (see Dunn and Dunn, 1981). The PPVT was designed for use with individuals aged 3 to 40 years. The English language version of the PPVT-R consists of 175 vocabulary items of generally increasing difficulty. The child listens to a word uttered by the interviewer and then selects one of four pictures that best describes the word's meaning. The PPVT-R has been administered, with some exceptions, to NLSY79 children between the ages of 3-18 years of age until 1994, when children 15 and older moved into the Young Adult survey. In the current survey round, the PPVT was administered to children aged 4-5 and 10-11 years of age, as well as to some children with no previous valid PPVT score.

  The first item, or starting point, is determined based on the child's PPVT age. Starting at an age-specific level of difficulty is intended to reduce the number of items that are too easy or too difficult, in order to minimize boredom or frustration. The suggested starting points for each age can be found in the PPVT manual (see Dunn and Dunn, 1981).

  Testing begins with the starting point and proceeds forward until the child makes an incorrect response. If the child has made 8 or more correct responses before the first error, a "basal" is established. The basal is defined as the last item in the highest series of 8 consecutive correct answers. Once the basal is established, testing proceeds forwards, until the child makes six errors in eight consecutive items. If, however, the child gives an incorrect response before 8 consecutive correct answers have been made, testing proceeds backwards, beginning at the item just before the starting point, until 8 consecutive correct responses have been made. If a child does not make eight consecutive responses even after administering all of the items, he or she is given a basal of one. If a child has more than one series of 8 consecutive correct answers, the highest basal is used to compute the raw score.

  A "ceiling" is established when a child incorrectly identifies six of eight consecutive items. The ceiling is defined as the last item in the lowest series of eight consecutive items with six incorrect responses. If more than one ceiling is identified, the lowest ceiling is used to compute the raw score. The assessment is complete once both a basal and a ceiling have been established. The ceiling is set to 175 if the child never makes six errors in eight consecutive items.

A child's raw score is the number of correct answers below the ceiling. Note that all answers below the highest basal are counted as correct, even if the child answered some of these items incorrectly. The raw score can be calculated by subtracting the number of errors between the highest basal and lowest ceiling from the item number of the lowest ceiling.

- **The Peabody Individual Achievement Test (PIAT): Math**

  The PIAT Mathematics assessment protocol used in the field is described in the documentation for the Child Supplement (available on the Questionnaires page). This subscale measures a child's attainment in mathematics as taught in mainstream education. It consists of 84 multiple-choice items of increasing difficulty. It begins with such early skills as recognizing numerals and progresses to measuring advanced concepts in geometry and trigonometry. The child looks at each problem on an easel page and then chooses an answer by pointing to or naming one of four answer options.

  Administration of this assessment is relatively straightforward. Children enter the assessment at an age-appropriate item (although this is not essential to the scoring) and establish a "basal" by attaining five consecutive correct responses. If no basal is achieved then a basal of "1" is assigned (see PPVT). A "ceiling" is reached when five of seven items are answered incorrectly. The non-normalized raw score is equivalent to the ceiling item minus the number of incorrect responses between the basal and the ceiling scores.

- **The Peabody Individual Achievement Test (PIAT): Reading Recognition**

  The Peabody Individual Achievement Test (PIAT) Reading Recognition subtest, one of five in the PIAT series, measures word recognition and pronunciation ability, essential components of reading achievement. Children read a word silently, then say it aloud. PIAT Reading Recognition contains 84 items, each with four options, which increase in difficulty from preschool to high school levels. Skills assessed include matching letters, naming names, and reading single words aloud.

  The only difference in the implementation procedures between the PIAT Mathematics and PIAT Reading Recognition assessments is that the entry point into the Reading Recognition assessment is based on the child's score in the Mathematics assessment, although entering at the correct point is not essential to the scoring.

The scoring decisions and procedures are identical to those described for the PIAT Mathematics assessment.

- **The Peabody Individual Achievement Test (PIAT): Reading Comprehension**

  The Peabody Individual Achievement Test (PIAT) Reading Comprehension subtest measures a child's ability to derive meaning from sentences that are read silently. For each of 66 items of increasing difficulty, the child silently reads a sentence once and then selects one of four pictures that best portrays the meaning of the sentence.

  Children who score less than 19 on Reading Recognition are assigned their Reading Recognition score as their Reading Comprehension score. If they score at least 19 on the Reading Recognition assessment, their Reading Recognition score determines the entry point to Reading Comprehension. Entering at the correct location is, however, not essential to the scoring.

  Basals and ceilings on PIAT Reading Comprehension and an overall nonnormed raw score are determined in a manner identical to the other PIAT procedures. The only difference is that children for whom a basal could not be computed (but who otherwise completed the comprehension assessment) are automatically assigned a basal of 19. Administration instructions can be found in the assessment section of the Child Supplement.

E Monte Carlo Exercise

We implement a Monte Carlo exercise to examine the properties of our estimator. The true data generating process is assumed to be the our estimated model, with some additional parametric assumptions about the measurement error process. In order to simulate the dataset, we use both the estimated measurement parameters and the joint distribution of children skills and investments. In addition, we assume that all the measurement noises are Normally distributed. We assume that the standard deviation of the error terms for all the skills measures are 0.5 (children and mothers) while we fix to 0.1 the standard deviation of the error terms for all the investment measures.

We generate a simulated longitudinal dataset of 10,000 children ranging from age 5-6 to age 13-14. In particular, the Monte Carlo analysis is performed estimating the model on 200 simulated data sets. In the following tables we show the mean estimates over the 200 estimates of the coefficients.

We focus only on estimates of skills technology, investment process and children's skills measurement parameters. Tables E-1-E-3 show true and mean estimated parameters. Overall, the estimator is able to recover the true parameters with minimal bias.

Table E-1: Monte Carlo Estimates for Investment Process

Parameter	True Parameters Age 5-6	Age 7-8	Age 9-10	Age 11-12	Monte Carlo Estimates Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.230	0.027	0.020	0.018	0.249	0.026	0.020	0.018
Log Mother Cognitive Skills	0.071	0.004	0.012	-0.005	0.077	0.002	0.008	-0.011
Log Mother Noncognitive Skills	0.359	0.742	0.694	0.712	0.322	0.748	0.700	0.700
Log Family Income	0.341	0.227	0.274	0.275	0.352	0.224	0.272	0.292
Variance Shocks	1.186	1.019	0.868	1.087	1.263	0.993	0.827	1.103

Notes: This table shows both the true estimates (reported also in Table B-2) and the mean Monte Carlo estimates for the investment equation. Each column shows the coefficients of the investment equation at the given ages. The dependent variable is investment in period t which is determined by the covariates at time t . For example, the first column shows the coefficients at age 5-6 for parental investments and child's skill and family income at age 5-6 as well.

Table E-2: Monte Carlo Estimates for Skill Technology

	True Parameters				Monte Carlo Estimates			
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	1.966	1.086	0.897	1.065	1.955	1.091	0.897	1.071
Log Investment	0.799	0.695	0.713	0.252	0.759	0.700	0.839	0.502
(Log Skills * Log Investment)	-0.105	-0.005	-0.003	0.003	-0.092	-0.005	-0.005	-0.002
Return to scale	2.660	1.776	1.606	1.320	2.623	1.786	1.731	1.571
Variance shocks	5.612	4.519	3.585	4.019	5.613	4.520	3.586	4.018
Log TFP	13.067	14.747	11.881	2.927	13.060	14.689	11.801	2.594

Notes: This table shows both the true estimates (reported also in Table 2) and the mean Monte Carlo estimates for the technology of skills formation. Each column shows the coefficients of the technology of skills formation at the given age. The dependent variable is log skills in the next period $t+1$ while the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8.

Table E-3: Monte Carlo Estimates for Measurement Parameters

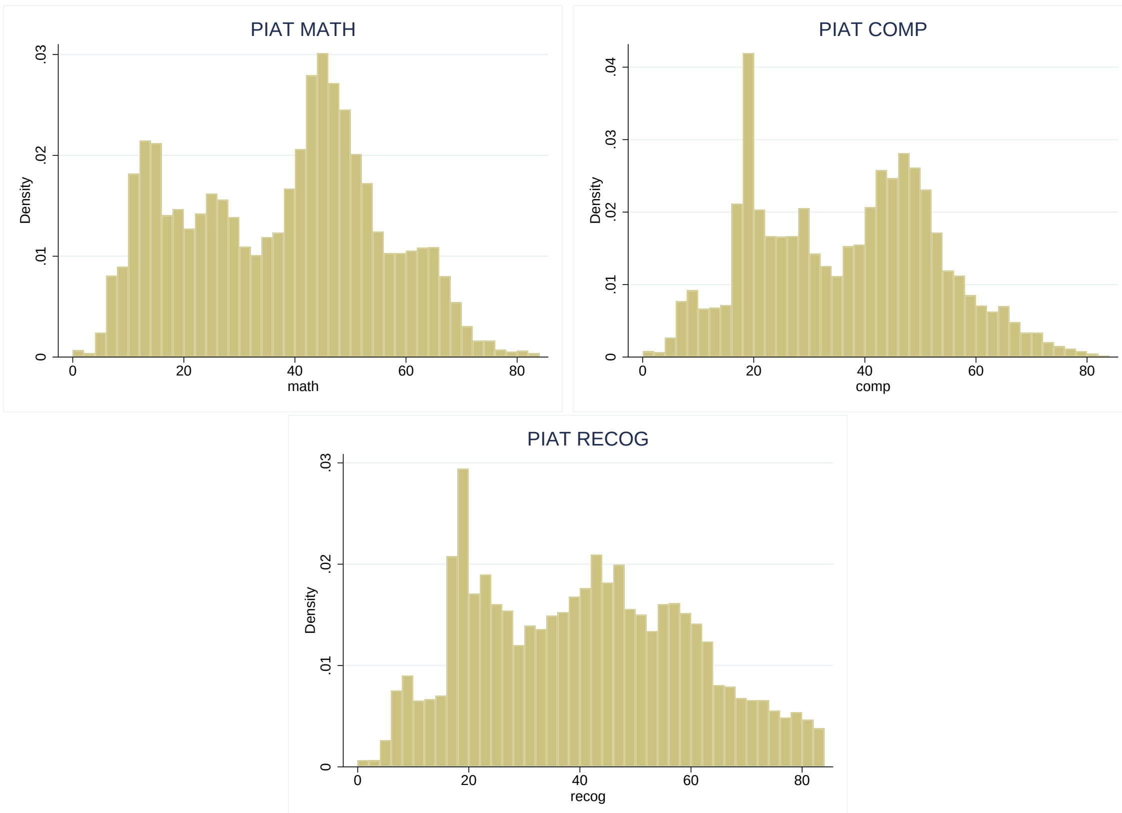
Parameter	True Constant (μ) Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14	Monte Carlo Constant (μ) Estimates Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14
PIAT Math	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858
PIAT Recognition	12.864	15.592	10.297	2.107	8.556	12.864	15.592	10.298	2.110	8.555
PIAT Comprehensive	12.770	15.014	12.273	6.129	9.041	12.770	15.013	12.270	6.132	9.040

Parameter	True Factor Loadings (λ) Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14	Monte Carlo Factor Loadings (λ) Estimates Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14
PIAT Math	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PIAT Recognition	2.238	0.906	1.136	1.347	1.195	2.238	0.906	1.136	1.347	1.196
PIAT Comprehensive	2.159	0.802	0.936	1.089	1.002	2.159	0.802	0.936	1.089	1.002

Notes: This table shows both the true estimates (reported also in Table B-11) and the mean Monte Carlo estimates for the measurement parameters of children skills measures equation. Each column shows the parameters at the given ages for each test score.

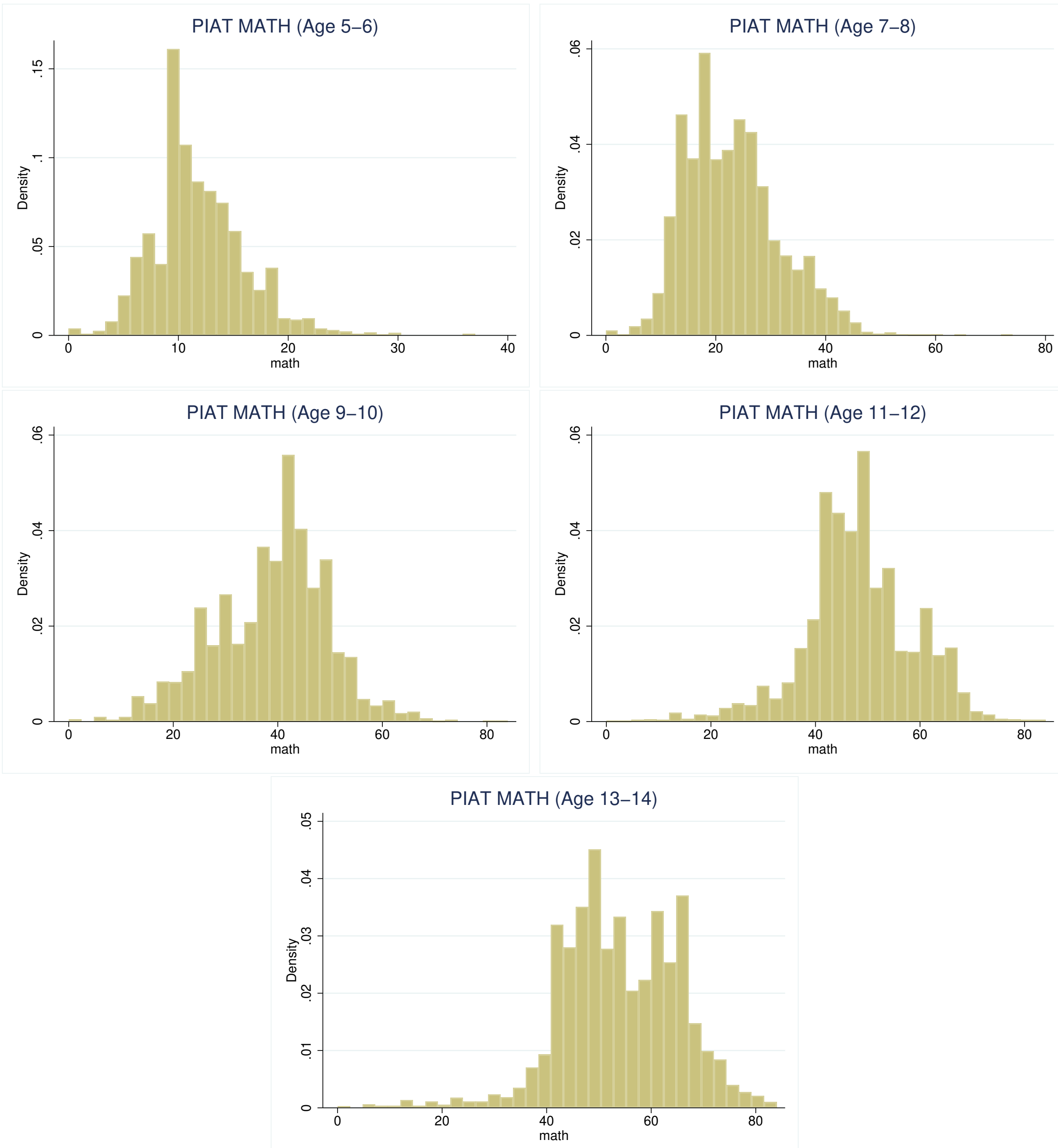
F Distribution of Test Scores

Figure F-1: Distribution of PIAT Math, Reading Comprehension, and Reading Recognition



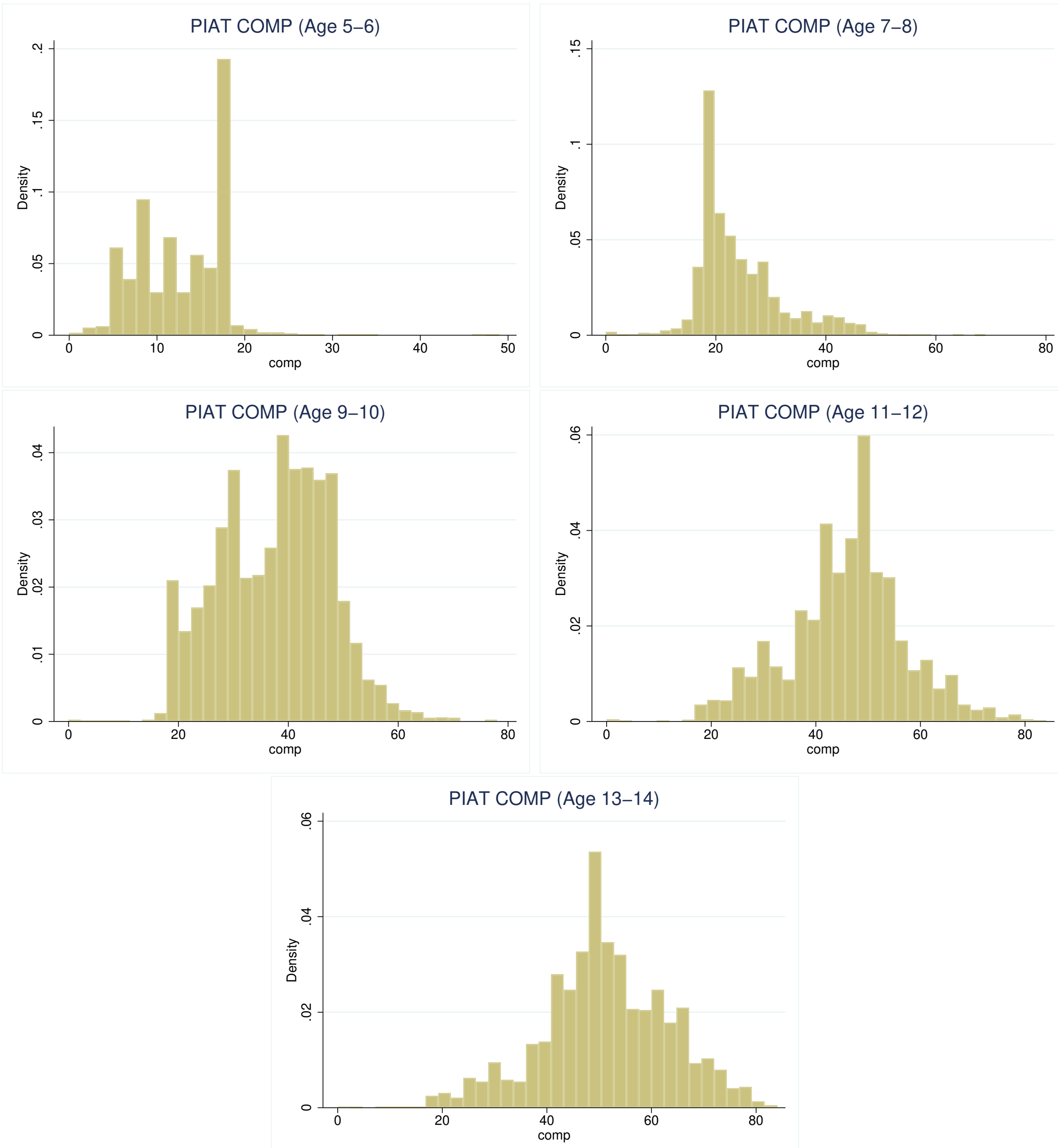
The figure shows the distribution in our sample of PIAT Math, PIAT Reading Comprehension, and PIAT Reading Recognition.

Figure F-2: PIAT Math by Children's Age



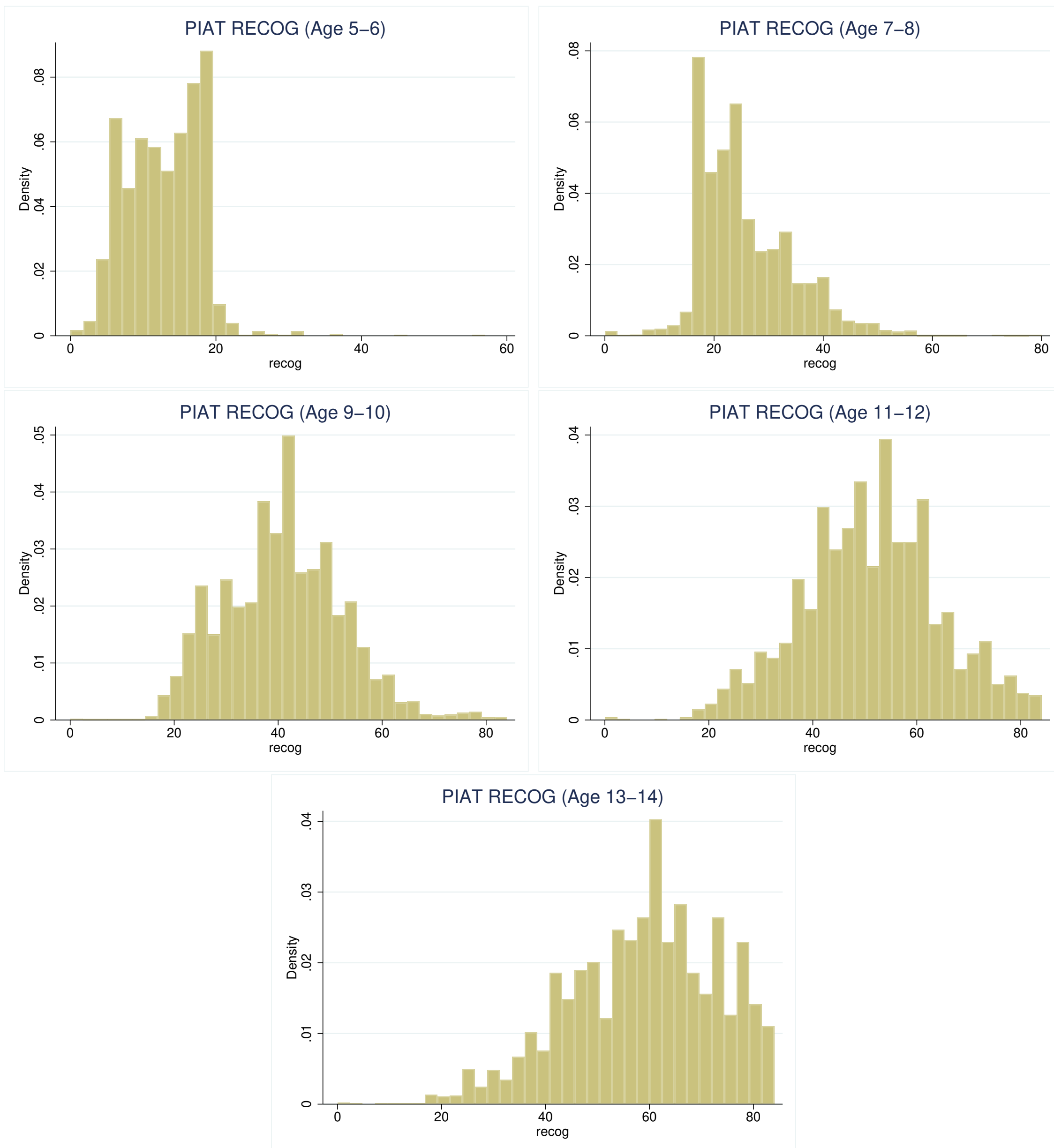
The figure shows the distribution in our sample of PIAT Math by children's age.

Figure F-3: PIAT Reading Comprehension by Children's Age



The figure shows the distribution in our sample of PIAT Reading Comprehension by children's age.

Figure F-4: PIAT Reading Recognition by Children's Age



The figure shows the distribution in our sample of the PIAT Reading Recognition by children's age.