



Childcare and parenting in the production of early life skills[☆]

Sebastián Gallegos^a, Jorge Luis García^{b,*}

^a Universidad Adolfo Ibáñez (UAI) Business School, Chile

^b Department of Economics, Texas A&M University, 2935 Research Parkway, College Station, TX 77845, USA

ARTICLE INFO

JEL classification:

C38

J13

J24

Keywords:

Childcare

Early life skills

Parental investment

Parenting

Skill formation

Skill production

ABSTRACT

We use data from a randomized early childhood education program to estimate the production technology of early life skills. Estimates indicate that, for more disadvantaged children, parental investment is a more effective input for producing skills than childcare. The reverse is true for the more advantaged. The program increases childcare for all children; it increases parental investment for the more disadvantaged. Therefore, our results indicate that programs stimulating parental investment promote mobility across the distribution of early life skills. We thus micro-found recent studies showing that successful early childhood education programs foster parental investment on top of offering high-quality childcare.

1. Introduction

An extensive literature documents the generality that high-quality early childhood education is an effective and socially efficient public policy, primarily when targeted to disadvantaged children.¹ This effectiveness and social efficiency are grounded in properties of the technology of skill formation. That is, they are grounded on properties of the mapping between the stock of skills a child has at a given period as a function of the stock of skills she had in the previous period and investments made in her. Specifically, early in life, her stock of skills and the investments made in her are direct substitutes or weak complements in the production of skills for the following period. The extent to which her stock of skills and investments made in her complement each other grows over time. Therefore, investing is most effective at remediating a low stock of skills earlier, as substitutability is likelier than in later periods. If remediation of a low stock of skills through investments occurs, follow-up investments become more effective as complementarity between future stocks of skills and investments, and, therefore, (dynamic) complementarity between investments at different periods arises.

This paper estimates a technology of skill formation or production function of skills early in life (up to age three). We consider one dimension of skills. This is standard practice when analyzing early life skills, as measuring skills multidimensionally at this early age is challenging. We allow for multidimensional investments. Generally, the empirical literature operationalizes the technology of skill formation by assuming that at any given period, there is one dimension of investment or, alternatively, that investment is adequately summarized with a single measure of time and material resources that parents spend on their children (henceforth, parenting). We allow for two dimensions of investment: parenting and time spent in high-quality center-based childcare (henceforth, childcare). This feature of our estimation allows us to conclude regarding the effectiveness of two inputs targeted by successful early childhood education programs.

We use data from the Infant Health and Development Program (IHDP). IHDP was a randomized early childhood education program designed after two iconic programs, the Carolina Abecedarian and Perry Preschool Projects, in eight sites located across the US. The data contain rich measures of cognition at age three, which we use to form the output of the technology of skill formation at age three (i.e., the stock

[☆] The Promising Practices Network gathered the data used in this paper, and data access was gained through the University of Michigan's Interuniversity Consortium for Political and Social Research data depository. Infant Health and Development Program, Phase IV, 2001–2004 U.S.; ICPSR #23580. Gallegos thanks the Center for Studies of Conflict and Social Cohesion (ANID/FONDAP/15130009 and ANID/FONDAP/1523A0005) and acknowledges financial support from the CONICYT FONDECYT (Grant numbers: 11220263) and CONICYT FONDAP (Grant numbers: 15130009, 1523A0005). The usual caveats apply.

* Corresponding author.

E-mail address: jlgarcia@tamu.edu (J.L. García).

¹ Surveys and studies documenting this fact include Cunha et al. (2006), Elango et al. (2016), and Heckman and Mosso (2014); and Bennis et al. (2024) and García et al. (2020).

of skills at age three); birth weight, gestational age, baseline maternal cognition, and child mental development at age one, which we use to form the stock of skills in the previous period (in our setting, the initial stock of skills); and several measures of parenting and childcare, which we use to measure the two dimensions of investment. Random assignment to the treatment group provided children and their families with home visits from birth to age three and high-quality center-based childcare at ages two and three. Control-group children did not receive home visits or high-quality childcare. While their parents were free to enroll them in other forms of childcare, their enrollment in substitutes for the high-quality childcare provided by assignment to treatment was minor.

We first estimate latent or factor scores of the stock of skills at age three, as well as the initial stock of skills and parenting, using measurement systems, following seminal work (Cunha & Heckman, 2008; Cunha et al., 2010) further explored in recent studies (Agostinelli & Wiswall, 2016a, 2016b). We then postulate a production function that allows the output-investment elasticities to vary as a function of the initial stock of skills. The production function also allows the two investment dimensions to interact. Identification of the production function relies on an instrumental-variable strategy that exploits random assignment to treatment and its interaction with baseline variables, potentially driving the effectiveness of childcare and parenting as inputs. Agostinelli and Wiswall (2016b) show that, in general, the units of the stocks of skills need to be comparable across periods for the elasticities to be interpretable. This requirement is not necessary in our setting, in which we estimate the technology at one period (age three), and the stock of skills in the previous period is an initial condition. In other settings, the technology of skill formation recurses, and the stock of skills appears as an input and an output. That is not our case, as data limitations only allow us to estimate the technology at one point in time.

Our findings indicate that childcare is most effective at producing early life skills for the most advantaged children in terms of initial skills: while the input elasticity does not differ statistically from 0 for the most disadvantaged (those below the 25th percentile of the distribution of initial skills), a 1% increase in childcare for those above the 75th percentile generates a statistically significant increase in early life skills of between 2.5% and 5%. In turn, parenting produces early life skills for the most disadvantaged most effectively. For them, a 1% increase in parenting generates a statistically significant increase in early life skills between 0.4% and 1.2%. For reference, assignment to the treatment group of IHDP increases childcare and parenting by 1.83 (s.e. 0.09) and 0.138 (s.e. 0.076) standard deviations, respectively.

A stylized counterfactual exercise allows us to clarify our findings and units of measurement. Assignment to the treatment group of IHDP increases childcare uniformly across the distribution of initial skills. Thus, boosting this input does not promote the mobility of the most disadvantaged children. Indeed, it exacerbates inequality because childcare is most effective for the most advantaged children. However, assignment to the treatment group increases parenting for the most disadvantaged. This input is the most effective for disadvantaged children. Holding all else equal, the increase in parenting generated by random assignment makes the average early life skills of the most disadvantaged almost identical to the observed average among the control (no-treatment) average. Put differently, boosting parenting for the most disadvantaged allows them to catch up to the average skills of all children, holding all else equal.

Our two main findings are (i) parenting is the most effective input in producing early life skills for the most disadvantaged, and (ii) high-quality early childhood education programs, like IHDP, increase this input the most for these children. These findings relate our work to a recently documented generality across high-quality early childhood education programs (García & Heckman, 2023). Namely, that programs effectively improving child skills and lifetime outcomes do so by encouraging the investment they receive from their parents, on top

of providing center-based childcare services. Therefore, we provide a micro-foundation for a recently documented generality: the effectiveness of high-quality early childhood education programs is based on the fact that they increase the input that happens to be the most effective at producing skills early in life, a sensitive period in which investments foster the productivity of subsequent investments.

A suggestive exercise using limited data at ages five and eight indicates that the patterns of the childcare and parenting elasticities with respect to initial skills persist, though they decrease in magnitude. This finding suggests that parental investments are more efficient earlier in life and that they are to be reinforced to further generate mobility of disadvantaged children across skill distributions, which, in turn, has been shown to promote social mobility in lifetime outcomes (e.g., Heckman et al., 2013).

2. Context and data

The Infant Health and Development Program (Ramey et al., 1992) aimed to foster the development of children at a socioeconomic disadvantage, measured by prematurity and low birth weight (Gross et al., 1997). It had a randomized design. The control and treatment groups of the program received pediatric follow-ups when children were between 40 weeks and 36 months old. The children in the treatment group received support from their parents through home visits and center-based childcare between childbirth and age three.

2.1. Treatment services

Home Visits. Professional home visitors visited the households of the treated children, training the parents in problem-solving and parenting skills, following the curriculum *Partners for Learning*. Professionals demonstrated, practiced, and discussed the curriculum with parents to train them as partners in their children's learning. Home visits occurred weekly during the first year of treatment and twice a month in the second and third years.²

High-Quality Childcare. Treatment at the childcare centers began when children reached age one and lasted two years. Children could spend between four and nine hours in childcare during weekdays, with the actual number of hours being chosen by their parents. The childcare centers, for the exclusive use of the treated children, satisfied state licensing requirements and other standard requirements to be considered high quality (e.g., low caregiver-child ratios). Caregivers were professionals; they continued *Partners for Learning* and introduced children to a set of activities based on *Early Partners*, a continuation of *Partners for Learning*.

2.2. Sample

Initial Sample. After an initial prescreening, 1028 mothers who gave birth to low-birthweight singleton children or twins between January and October of 1985 in eight university hospitals (henceforth, sites) across the United States agreed to participate in the program.³ The randomization was stratified by site, treatment status, and birth weight—low-low (≤ 2000 g) or low-high (> 2000 g, ≤ 2500 g). The per-stratum probability of being treated was 1/3. After randomization, 43 mothers withdrew from the program. This source of attrition is minor,

² Home visits are a common component of high-quality early childhood education programs. The Perry Preschool Project and Head Start are examples. These visits are an effective component of the Head Start curriculum (Walters, 2015).

³ The sites were in Arkansas, Connecticut, Florida, Massachusetts, New York, Pennsylvania, Texas, and Washington.

and we do not address it. The 985 remaining mothers gave birth to the children in the initial sample.

Main Analysis Sample. The initial sample contained 882 mothers of singletons (547 controls and 335 treatments) and 103 mothers of twins (61 controls and 42 treatments). The IHDP oversampled twins: 2% of children born in the US in 1985 were twins (National Bureau of Economic Research, 2022), while 10% of children in the initial sample were twins. This oversampling occurs because twins are eight times more likely to be premature and have low birth weight than singletons (National Center for Health Statistics, 2022). Unlike singletons, twins are likely to be premature and have low birth weight for biological reasons, holding all else equal.⁴ Dougan et al. (2023) argue that the data-generating processes for singletons and twins differ in economically meaningful ways (e.g., there are increasing returns in parenting twins). They show that the treatment effects of the IHDP on singletons resemble the treatment effects of programs like the Carolina Abecedarian and Perry Preschool Projects. In contrast, the treatment effects on twins are minor. Therefore, pooling singletons and twins attenuates the treatment effects of IHDP, obscuring the conclusions regarding the treatment effects on most of the sample. For clarity, we follow standard practice in studies of IHDP (e.g., Chaparro et al., 2020) and do not consider twins when forming the main analysis sample, noting that treatment-control average differences at baseline remain balanced after dropping the twins (see Panels a. and b. of Table 1).

The skills, childcare, and parenting measures used to construct the variables described below are based on comprehensive questionnaires. Naturally, these variables have item non-response. Therefore, the number of singleton children in the main analysis sample decreases from 882 in the initial sample to 644. Yet, standard baseline-covariate tests in Panels a. and b. of Table 1 indicate covariate balance between the treatment and control groups. In addition, (i) our estimators account for baseline covariates that could help account for non-random item non-response; and (ii) Dougan et al. (2023) document that item non-response is indeed random. Item non-response is a minor issue that we do not address.⁵

2.3. Measures of skills, childcare, and parenting

Our main results are based on estimating the production technology of early life skills at age three. The three inputs we consider are initial skills, childcare, and parenting. Data limitations do not allow us to consider periods between our observation of initial skills at baseline and age three. In practice, there is no recursion in the production technology that we estimate. Initial skills are simply an initial condition, and differences in units between initial and early life skills do not contaminate the interpretation of the estimated parameters. If the production technology recursed, skills would appear as an output in one period and as an input in subsequent periods, making unit comparability essential for interpreting parameter estimates (Agostinelli & Wiswall, 2016b; Freyberger, 2022).

Skills are unobserved. Therefore, the usual practice is to follow seminal work (i.e., Cunha & Heckman, 2008; Cunha et al., 2010) and estimate latents or factor scores using multiple imperfect or noisy skill measures. An alternative is to estimate parameters determining the distribution of these latents together with the parameters characterizing the production technology (e.g., Agostinelli & Wiswall, 2016a; Cunha et al., 2010). This strategy is especially suitable when the production technology is recursive. In our case, there are no gains from joint

estimation other than statistical efficiency. Separate estimation allows us to clearly describe the latents and associated raw, noisy measures.

Initial Skills. In principle, there are several skills at any given age. Measuring multiple skills is challenging when children are recently born or early in their lives. We estimate a single latent or factor score and interpret it as a general descriptor of a child's capacity to function using a basic framework (e.g., Williams, 2020). Identification requires at least three noisy measures, all linear functions of the underlying latent and measurement error. Measurement error is assumed to be mean-independent of the underlying factor and independent across measurements.

We estimate the latent of initial skills using four measures: birth weight, gestational age, the Bayley Mental Development Index (MDI) at baseline,⁶ and a measure of maternal intellectual quotient (IQ) based on the Peabody Picture Vocabulary Test.⁷ The first two are typical measures in studies of early life skills and human capital (e.g., Aizer & Cunha, 2012; Rosenzweig & Wolpin, 1988).⁸ The third is a more direct measure of child functioning. We also include maternal IQ at baseline, naturally correlated with initial child skills, for a more precise estimation. Panel (a) of Fig. 1 displays the distribution of the latent initial skills by treatment status, and Panels a. and b. of Table 1 display the associated measures; they verify no average treatment-control difference, an essential requirement for these baseline measures.

Early Life Skills. Similar to initial skills, we estimate a latent summarizing early life skills. In this case, the latent is based on the three measures summarized in Panel d. of Table 1. These three measures are standardized to the national mean of 100 and standard deviation of 15. They show that, between ages two and three, assignment to the IHDP treatment generates an average increase of between one-third and almost two-thirds of the standard deviation of the national distribution of the respective measure.

Latents do not have a natural location or scale, so it is innocuous to standardize them arbitrarily. We standardize the initial and early life skills latents by subtracting from them the control-group mean, dividing them by the control-group standard deviation, and adding to them 100. This standardization allows us to interpret treatment-control differences straightforwardly. For instance, assignment to the IHDP treatment generates an average increase of 0.42 (s.e. 0.07) in early life skills. This impact equals 42% of the no-treatment counterfactual standard deviation; equivalently, it equals a 0.42% increase with respect to the control-group mean. We standardize the variables so that a 1 standard deviation average increase is equivalent to a 1% average increase from the control group.

Childcare. Mothers of all children in the main analysis sample reported the weekly hours their children spent in childcare centers in an average week. The reports occurred when the children were 18, 24, 30, and 36 months old. To consolidate this information into one measure, we average observations of the average week across these four ages. This

⁴ The oversampling of twins was unlikely due to in-vitro fertilization methods, which were incipient in the early 1980s (Wang & Sauer, 2006).

⁵ Table A.1 indicates that, in terms of the observed average for the treatment and control groups, the sample of 882 for whom we fully observe basic baseline characteristics and the main analysis sample of 644 children are virtually identical.

⁶ The Bayley Mental Development Index was measured throughout age one; it was not strictly a baseline variable. However, Table 1 indicates that the treatment had not generated an impact. Using this as a baseline measure is justified and helpful, as it is a powerful indicator of a child's capacity to function. Similarly, for early life skills, we combine measures at age three with a measure at age two and interpret the resulting latent as an age-three latent. The age-two measure makes identification possible.

⁷ For estimation, we use an asymptotic distribution-free estimator that efficiently weights covariance moments across the four measures (Browne, 1984).

⁸ Birth weight is especially useful in this setting, where all children are low birth weight. All observed children have a birth weight that is considered unhealthy. Therefore, a marginal increase would put a child in our sample at a relative advantage in terms of capacity to function with respect to the other children (Boardman et al., 2002; Boulet et al., 2011).

Table 1
Summary statistics: Main analysis sample.

	(1)	(2)	(3)	(4)
	Control (mean)	Treatment (mean)	Difference (1) – (2)	Standard error (of difference)
Panel a. Child at Baseline				
Male	0.487	0.481	–0.006	0.040
Black	0.510	0.516	0.005	0.040
Hispanic	0.083	0.093	0.010	0.023
Gestational Age (weeks)	33.047	33.140	0.093	0.211
Birth Weight (kilograms)	1.798	1.838	0.040	0.036
Bayley Mental Development Index	111.269	110.612	–0.657	1.234
Panel b. Mother and Household at Baseline				
Age	24.912	24.562	–0.350	0.483
IQ-Peabody Picture Vocabulary Test	82.231	82.101	–0.130	1.661
Education	12.681	12.287	–0.395	0.196
Works	0.396	0.376	–0.020	0.039
Married	0.513	0.442	–0.071	0.040
Other Children in Household	0.687	0.702	0.015	0.080
Employed Adults in Household	1.111	1.012	–0.100	0.060
Panel c. Childcare and Parenting (Ages 1 to 3)				
Childcare Hours per Week	5.571	26.971	21.400	1.042
Parenting	100.000	100.138	0.138	0.076
Panel d. Measures of Early Life Skills (Ages 2 to 3)				
Bayley Mental Development Index	96.759	106.384	9.625	1.452
IQ-Peabody Picture Vocabulary Test	85.528	93.636	8.107	1.419
IQ-Stanford-Binet Test	85.915	92.221	6.306	1.389

Note: Columns (1) and (2) display the average of the variable indicated in the row for the control and treatment groups. Column (3) displays the difference between Columns (2) and (1), and Column (4) displays the robust standard error of this difference clustered at the child level. All computations are based on the main analysis sample of 644 singleton participant children (observed with complete item response up to age three).

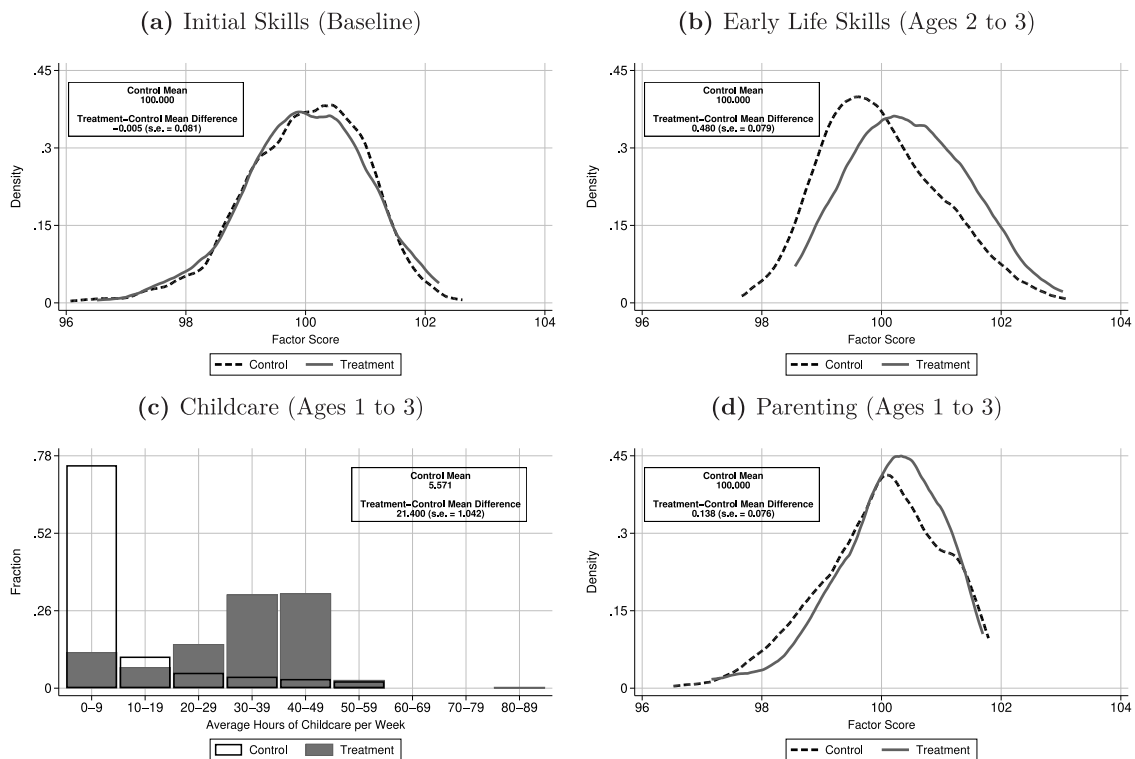


Fig. 1. Measure description: Initial skills, early life skills, and inputs.

Note: Panel (a) displays the empirical distribution of initial skills at baseline (a factor score of the child gestational age, birth weight, Bayley Mental Development Index at age one and the mother Peabody Picture Vocabulary Test) by treatment status. Panel (b) is analogous in format to Panel (a) for early life skills (a factor score of the child Bayley Mental Development Index at age two, Peabody Picture Vocabulary Test at age three, and Stanford-Binet Test at age three). Panel (c) displays a histogram of the childcare hours per week by treatment status (childcare hours per week is computed as the average of the childcare hours per week during a regular week when the child is 12, 18, 24, 30, and 36 months old). Panel (d) is analogous in format to Panel (a) for parenting. Parenting is the average of two factor scores based on the Home Observation Measurement of the Environment at ages 1 and 3. All panels display the control-group mean and the average treatment-control difference, as well as the robust standard error of this difference clustered at the child level. The latents or factor scores in Panels (a), (b), and (d) are standardized by subtracting from them the control-group mean, dividing them by the control-group standard deviation, and adding to them 100.

average is our measure of the “childcare” or “childcare between ages one and three” input. We do not correct it for measurement error.

Mothers did not report the specific center that their children attended. However, program records indicate that treatment-group children exclusively attended IHDP childcare centers (Gross et al., 1997). We assume that the childcare hours mothers reported were spent at the IHDP childcare centers, which were of homogeneous, high quality. Control-group children had no access to the IHDP childcare centers. However, their mothers were free to enroll them in alternative centers of potentially varying quality. This enrollment was relatively minor: an average of 5.6 h a week, which was, on average, 21.4 (s.e. 1.04) fewer hours than the treatment group. Therefore, we consider the lack of quality adjustment in their enrollment hours a minor issue of our childcare measure (indeed, 73% control-group children did not attend childcare at all; for them and the entire treatment group, there is no measurement issue at all).

For the analysis below, we standardize the variable childcare by subtracting the control-group mean from it, dividing it by the control-group standard deviation, and, finally, relocating it by adding the control-group mean. This rescaling is such that all values are positive, and taking logs is not problematic when estimating the production function. This type of rescaling is suggested when the dependent variable is truncated at 0 (Chen & Roth, 2024). The standardized childcare variable is described in Fig. A.1. It has a control-group mean of 5.6 and an average treatment-control difference of 1.83 (s.e. 0.09).

Parenting. Like skills, parenting is not observed. We estimate two latents, one at age one and another at age three, using the long form of the Home Observation Measurement of the Environment (HOME) inventory (Bradley & Caldwell, 1984; Bradley et al., 1992), which was used to measure the interaction of the child and mother participants, as well as the resources available during that interaction.

For age one, the HOME inventory has 24 binary items (e.g., parent structures child’s play periods) divided into six subscales. Subscale scores are the fraction of items within the subscale for which the binary items are one. We use the scores in the six subscales as measures of parenting at age one to estimate a latent, using the same procedure described for constructing the latents of initial and early life skills. We proceed in the same way to estimate the latent at age three.⁹ Factor analyzing the sub-scales of the HOME scores produces a data-driven, error-corrected measure of the time and material goods spent by parents on children, which we call “parenting”. Below, we use the average of the two latents as our measure of parenting, which we standardize in the same way that we standardize the latents of skills. Panel (d) of Fig. 1 shows that assignment to treatment substantially shapes parenting.

3. Framework and empirical strategy

The effectiveness and social efficiency of investing in high-quality early childhood education are founded in properties of the function f_t in

$$\underbrace{\theta_{t+1}}_{\text{skills at time } t+1} = f_t \left(\underbrace{\theta_t}_{\text{stock of skills at } t}, \underbrace{I_t}_{\text{investments at } t} \right). \quad (1)$$

The literature finds that early in life, skills (θ_t) and investments (I_t) at any given period t are direct substitutes or weak complements in the production of skills at $t+1$ (θ_{t+1}) and that complementarity

⁹ At age one, the subscales of the HOME inventory are parental warmth, parental verbal skills, parental lack of hostility, learning and literacy, activities and outings, and developmental advance. At age three, the HOME inventory has eight subscales (learning stimulation, access to reading, parental verbal skills, parental warmth, home exterior, home interior, outings and activities, and parental lack of hostility).

between skills and investments grows over time. Therefore, earlier investments I_t are most effective at remediating relatively low levels of skills θ_t , because substitutability is likelier than in later periods. When investment remediates low stocks of skills, follow-up investments become more effective as complementarity between future stocks of skills and investments, and, therefore, (dynamic) complementarity between investments at different periods arises.

We aim to estimate f_t at age three, considering our age-three measure of early life skills as an output. In our case, θ_{t+1} , θ_t , and f_t scalar. Though the empirical literature operationalizes the technology in Eq. (1) by assuming that at any given period t , I_t has one dimension, we consider two dimensions (childcare and parenting). We estimate latents or factor scores of θ_{t+1} , θ_t , and parenting as described in Section 2.3. As described in that section, a measure of childcare is readily available.¹⁰

We use a transcendental logarithmic (translog) production function as the empirical counterpart of Eq. (1):

$$\begin{aligned} \ln(\theta_{\text{early}}) = & \alpha_\theta \cdot \ln(\theta_{\text{initial}}) \\ & + \alpha_c \cdot \ln(\text{childcare}) + \alpha_{c^2} \cdot [\ln(\text{childcare})]^2 \\ & + \alpha_{c\theta} \cdot \ln(\text{childcare}) \cdot \ln(\theta_{\text{initial}}) \\ & + \alpha_p \cdot \ln(\text{parenting}) + \alpha_{p^2} \cdot [\ln(\text{parenting})]^2 \\ & + \alpha_{p\theta} \cdot \ln(\text{parenting}) \cdot \ln(\theta_{\text{initial}}) \\ & + \alpha_{cp} \cdot \ln(\text{childcare}) \cdot \ln(\text{parenting}) \\ & + \varepsilon, \end{aligned} \quad (2)$$

where θ_{early} , θ_{initial} , childcare, and parenting are described in Fig. 1, and ε is an error term. This framework is cross-sectional: θ_{initial} is an initial condition at age three when θ_{early} is measured. The inputs parenting and childcare are measured between ages one and three.

Though our framework is cross-sectional, the production function we estimate accommodates non-linearities in the elasticities of the inputs with respect to initial skills, speaking to complementarity and substitutability patterns. Differently from Cobb–Douglas cases, the technology we consider does not restrict the elasticities of substitution between inputs, which have been found to differ from 1 (e.g., Cunha et al., 2010). This generality makes translog and constant-elasticity-of-substitution functions common skill-formation technologies (e.g., Agostinelli & Wiswall, 2016a; Cunha et al., 2010; Del Bono et al., 2022). A general practice in the literature is to estimate production technologies of various functional forms using different measures of parental investment or parenting as the only input. An exception is Attanasio et al. (2020a), who consider two forms of parental investment as inputs, i.e., material investment and time investment. A primary difference between our framework and others is that we consider an additional input, childcare, and its interactions with initial skills and parenting.

The parameters of interest are the elasticities of childcare and parenting with respect to θ_{early} , for which we need consistent estimates of the α parameters in Eq. (2). If ε is mean-independent of the right-hand-side variables of this equation, ordinary least squares (OLS) yields such estimates. This assumption is implausible because parents base their childcare and investment decisions on unobserved information in ε . Three alternatives are standard in the literature: (i) decomposing ε into a time-invariant component and a time-varying random component in a longitudinal setting, estimating the time-invariant component; (ii) exploiting sequential mean independence and lagged skills as instruments

¹⁰ As noted in the introduction, Agostinelli and Wiswall (2016b) show that, in general, the units of θ_{t+1} and θ_t need to be comparable for the parameters involving θ_t to be interpretable in terms of elasticities with respect to θ_{t+1} . This requirement is not necessary in our setting, in which we estimate the technology at one period, and, therefore, θ_t is an initial condition. In other settings, a recursion of f_t is considered, and θ_{t+1} appears as an input after being an output. That is not our case, as data limitations only allow us to estimate the technology at one point in time.

in a longitudinal setting; and (iii) exploiting control functions based on input prices in a cross-sectional setting.¹¹ We cannot implement the first two approaches in our setting. The third approach could be an option, but price variation is likely helpful in settings with a wider geography. Instead, we exploit randomization in an instrumental-variable (IV) setting.

3.1. Instrumental-variable strategy

The seven variables involving either childcare or parenting in Eq. (2) potentially fail to satisfy mean independence. Therefore, we need at least seven instruments.¹² The first instrument is randomization (R) to the treatment or control groups. We generate additional instruments by interacting R with a subset of baseline variables X . Importantly, all of our estimations of Eq. (2) control for baseline variables X , which includes θ_{initial} henceforth. Instruments based on interactions are described in Donald and Newey (2001) and applied in Angrist and Krueger (1991). Initial skills, the maternal characteristics in Panel b. of Table 1, and site fixed effects are likely to induce variation in treatment effectiveness. Interacting them with R likely leads to instruments with strong first stages.¹³ We interact these variables with R to generate thirteen instruments in addition to R .

We need at least seven instruments and have a total of fourteen. Newey (1990) proposes using the projection of each endogenous variable on the instruments and controls, generating efficient instruments.¹⁴ We form these projections and use them as instruments. Our estimation procedure consists of three steps: (i) projecting our seven potentially endogenous variables on the fourteen instruments and X ; (ii) use the projections obtained in (i) to run the standard first stage of linear IV estimation; and (iii) use the projections in (ii) to run the standard second stage of linear IV estimation.

Exogeneity. To be valid, our instruments must be exogenous (randomly assigned and excluded). We justify the exogeneity of R and the interaction instruments, which should translate into the exogeneity of the projection instruments we generate to gain efficiency.

By design, R is randomly assigned. Table 1 documents the resulting balance in observed characteristics between the treatment and control groups. We argue that R also satisfies the exclusion restriction because the literature indicates that, while assignment to high-quality early childhood education shifts the inputs of the production function of early life skills, it does not shift the parameters (Attanasio et al., 2020a; Heckman et al., 2013). Therefore, we assume that R is exogenous.

In our context, $\mathbb{E}[\epsilon|X] = 0$ is a plausible assumption because X only includes baseline variables. If this assumption holds, the interaction of R with any element of X is also a plausibly exogenous instrument. Of course, these interaction instruments are only valid conditional on X (e.g., the interaction of R with initial skills is only a plausibly exogenous instrument conditional on initial skills). A balance check of random assignment is the following. Let \tilde{x} denote a baseline variable in X that is not interacted to form instruments and \tilde{X} the subset of variables in X that are interacted to form instruments. We estimate

$$\tilde{x} = \beta \cdot \text{instrument} + \gamma \cdot \tilde{X} + v, \quad (3)$$

¹¹ Examples include (i) Cunha et al. (2010) and Pavan (2016); (ii) Del Bono et al. (2022), Agostinelli and Wiswall (2016a), and Cunha et al. (2010); and (iii) Attanasio et al. (2020b) and Attanasio et al. (2020a).

¹² As a measurement-error corrected predetermined stock, we assume that θ_{initial} alone does not generate endogeneity issues.

¹³ Elango et al. (2016) and Heckman and Mosso (2014) argue that initial endowments and paternal characteristics shape how early childhood education programs affect children. Bernal and Keane (2011) exploit maternal factors, such as marital status and state location, to derive instruments that shift childcare and employment decisions. Similar arguments motivate our interaction instruments.

¹⁴ This procedure reduces the number of instruments, avoiding the “many-instruments-problems” that biases IV estimators (Stock & Yogo, 2005).

Table 2

Instrument diagnostics: Additional randomization tests.

	(1)	(2)	(3)	(4)
	Male <i>t</i> -stat [<i>p</i> -value]	Black <i>t</i> -stat [<i>p</i> -value]	Hispanic <i>t</i> -stat [<i>p</i> -value]	Joint <i>F</i> -stat [<i>p</i> -value]
Randomization (R)	0.000 [0.500]	0.373 [0.355]	0.084 [0.467]	0.225 [0.879]
$R \times \text{Initial Skills}$	0.000 [0.500]	0.363 [0.358]	0.076 [0.470]	0.223 [0.881]
$R \times \text{Maternal Age}$	0.027 [0.489]	0.436 [0.332]	0.029 [0.489]	0.291 [0.832]
$R \times \text{Maternal IQ}$	0.076 [0.470]	0.020 [0.492]	0.000 [0.500]	0.153 [0.928]
$R \times \text{Maternal Education}$	0.041 [0.484]	0.092 [0.463]	0.011 [0.495]	0.185 [0.906]
$R \times \text{Mother Works}$	3.343 [0.000]	0.000 [0.500]	0.079 [0.468]	0.726 [0.537]
$R \times \text{Married Mother}$	4.500 [0.000]	0.061 [0.476]	0.000 [0.500]	0.758 [0.518]
$R \times \text{Site 1}$	0.023 [0.491]	0.039 [0.484]	0.239 [0.405]	0.360 [0.782]
$R \times \text{Site 2}$	0.000 [0.500]	0.000 [0.500]	0.040 [0.484]	0.180 [0.910]
$R \times \text{Site 3}$	0.023 [0.491]	0.000 [0.500]	0.201 [0.420]	0.190 [0.903]
$R \times \text{Site 4}$	0.001 [0.500]	0.308 [0.379]	0.355 [0.361]	0.858 [0.463]
$R \times \text{Site 5}$	0.007 [0.497]	0.000 [0.500]	4.218 [0.000]	0.703 [0.551]
$R \times \text{Site 6}$	0.000 [0.500]	0.000 [0.500]	0.000 [0.500]	0.013 [0.998]
$R \times \text{Site 7}$	0.014 [0.494]	0.000 [0.500]	0.011 [0.496]	0.085 [0.968]
$R \times \text{Site 8}$	0.575 [0.283]	36.379 [0.000]	1.185 [0.118]	2.436 [0.064]

Note: The first entry of Column (1) displays the *t*-statistic and corresponding (two-sided) *p*-value of the coefficient on treatment-control status or randomization indicator (R) in a regression of a male indicator on R . The regression also conditions on the initial skills factor score described in Panel (a) of Fig. 1, the maternal and household characteristics at baseline in Panel b. of Table 1, and site indicators. The first entries of Columns (2) and (3) are analogous in format to the first entry of Column (1) for indicators of black and Hispanic. The first entry of Column (4) displays the *F*-statistic and corresponding *p*-value of a joint test of significance of male, black, and Hispanic in a regression of R on these three variables and the initial skills factor score described in Panel (a) of Fig. 1, the mother and household characteristics at baseline in Panel b. of Table 1, and site indicators. The remaining rows in the table are analogous in format to the first row for each indicated variable (regressions use the interaction of R with the observed characteristic in the label instead of R). The *t*-statistics and *p*-values are based on robust standard errors clustered at the child level.

where β is a coefficient, γ is a coefficient vector, and v is an error term. Testing the null hypothesis $\beta = 0$ is a test of balance. If instrument = R , the test is a standard balance test conditional on \tilde{X} . We display this test for each of our instruments in Table 2 for $\tilde{x} = \{\text{Male, Black, Hispanic}\}$, which are variables in X that we do not interact to form instruments. We also display the corresponding joint *F*-test.¹⁵ The resulting tests largely indicate that the instruments satisfy random assignment.

As for the exclusion restriction, if R does not shift the production function, it is likely that its interaction with baseline characteristics does not shift the production function either. Therefore, the fourteen instruments, and, thus, the projection instruments, are plausibly exogenous.

¹⁵ That is, the *F*-test contrasting a model regressing the instrument on Male, Black, Hispanic, and \tilde{X} with a model regressing the instrument on \tilde{X} .

Rank. We estimate the first stage using the projection instruments. We directly test their rank. If there were one potentially endogenous variable and one projection instrument, the rank condition would be tested as follows: (i) regress the endogenous variable on the projection instrument and X , and (ii) contrast the regression model in (i) with a model excluding the projection instrument. The projection instrument would satisfy the rank condition if the F -statistic from the test in (ii) exceeded 10 (Stock & Yogo, 2005). When there are multiple potentially endogenous variables, it is possible to compute an analogous F -statistic: (i) regress the endogenous variable on its corresponding projection instrument and X , and (ii) contrast the model in (i) with a model excluding the projection instrument. Panel a. of Table 3 indicates strong first stages based on this test. However, the correlation between the endogenous variables could make the F -statistics overstate the rank of the projections instruments. We pursue formal, joint rank tests.

In the case of multiple endogenous variables, there is a system of first-stage equations and a matrix of coefficients associated with the projection instruments. The alternative hypothesis associated with a satisfactory rank is that the rank of this matrix equals the number of endogenous variables (in our case, 7). Null hypotheses of ranks below seven are tested against this alternative. There are χ^2 -statistics associated with each of these tests (Kleibergen & Paap, 2006). We display them in Panel b. of Table 3. We consider three different control sets for the individual and joint tests. The most comprehensive set includes initial skills, male, black, and Hispanic indicators, the maternal and household characteristics in Panel b. of Table 1, and site indicators. In all instances, we reject the null hypothesis that the rank is less than 6 (or any integer between 0 and 5, individually) with 1% significance. Our instruments plausibly satisfy the rank condition.

3.2. Inference

In Tables 1 to 3 and Panels (a) and (b) of Fig. 3, our inference is robust and clustered at the child level. Robust inference is standard. Clustering at the level of randomized assignment is recommended (Abadie et al., 2022). We use this inferential procedure for basic descriptions and tests because it is transparent and straightforward to replicate. For Table 4 and Figs. 2 and 4, our inference is bootstrapped, clustering at the child level and stratifying by site, treatment status, and birthweight group (low-low or high-low). Bootstrapping allows us to account for the sampling variation in each of the steps of our estimation (i.e., estimation of latents of skills and parenting, projections for constructing efficient instruments, and main estimation). Stratifying allows us to emulate the randomization procedure in each bootstrap resampling.

4. Results

Table 4 displays our main elasticity estimates evaluated at the mean. As a reference, Panel a. displays elasticities based on OLS estimates of Eq. (2). Panel b. displays elasticities based on IV, our preferred empirical strategy. In either case, estimates are robust to accounting for our control sets. While the IV estimates are naturally more imprecise, the estimates indicate that the elasticities differ statistically from 0 when using standard significance levels. We interpret results using the broadest control set (Control Set 3).

Initial skills are a baseline, measurement-error-corrected variable. The corresponding elasticity is unlikely to be biased, as suggested by the similarity between the OLS and IV elasticities of 0.26 (s.e. 0.025) and 0.19 (s.e. 0.055). Both estimates indicate that skills are self-productive, a necessary condition for dynamic complementarity. As we explain above, given that initial skills are an initial condition and our estimated production technology does not recurse, the difference in the measures on which initial and early life skills are based is not an issue. However, it is expected for the skill-to-skill elasticities

Table 3
Instrument diagnostics: Rank tests.

	(1)	(2)	(3)
	Control Set 1	Control Set 2	Control Set 3
Panel a. Instrument Specific Tests (Projection Instruments)			
F -stat			
[p -value]			
Childcare	518.520 [0.000]	529.997 [0.000]	586.140 [0.000]
Parenting	109.790 [0.000]	69.318 [0.000]	34.899 [0.000]
Childcare ²	517.039 [0.000]	528.490 [0.000]	587.258 [0.000]
Parenting ²	487.272 [0.000]	500.048 [0.000]	554.750 [0.000]
Childcare \times Parenting	109.991 [0.000]	69.445 [0.000]	34.906 [0.000]
Childcare \times Initial Skills	518.417 [0.000]	529.858 [0.000]	586.056 [0.000]
Parenting \times Initial Skills	109.979 [0.000]	69.430 [0.000]	34.841 [0.000]
Panel b. Joint Tests (All Projections)			
χ^2 -stat			
[p -value]			
H_0 : rank = 0; H_1 : rank = 7	305.944 [0.000]	299.838 [0.000]	285.427 [0.000]
H_0 : rank = 1; H_1 : rank = 7	122.596 [0.000]	119.088 [0.000]	111.754 [0.000]
H_0 : rank = 2; H_1 : rank = 7	80.698 [0.000]	78.024 [0.000]	62.169 [0.000]
H_0 : rank = 3; H_1 : rank = 7	44.778 [0.000]	40.945 [0.001]	35.731 [0.003]
H_0 : rank = 4; H_1 : rank = 7	18.442 [0.030]	18.636 [0.028]	20.342 [0.016]
H_0 : rank = 5; H_1 : rank = 7	8.584 [0.072]	8.409 [0.078]	14.371 [0.006]
H_0 : rank = 6; H_1 : rank = 7	4.065 [0.044]	3.850 [0.050]	7.145 [0.008]

Note: The entry Childcare in Column (1) of Panel a. displays the rank F -statistic and corresponding p -value of the projection instrument childcare. The statistic corresponds to a test contrasting a model that regresses childcare on the projection instrument and Control Set 1 and a model that regresses childcare on Control Set 1. The projection instrument childcare is the prediction of a regression of childcare on the instruments (a treatment-status indicator R and the interactions of R and a subset of the baseline variables listed in Table 2), and Control Set 1. The remaining entries in Column (1) of Panel a. are analogous in format to the entry Childcare for the other projection instruments. Column (1) of Panel b. displays the joint rank χ^2 -stats and corresponding p -values of the instruments in Panel a. The statistics correspond to tests of each of the indicated null and alternative hypotheses. The rank tests are conditional Control Set 1. Columns (2) and (3) of Panels a. and b. are analogous in format to Column (1) for Control Sets 2 and 3. Control Set 1: initial skills and site indicators. Control Set 2: Control Set 1 and male, black, and Hispanic indicators. Control Set 3: Control Set 2 and the maternal and household characteristics in Panel b. of Table 1. The statistics and p -values are based on robust standard errors clustered at the child level.

to be farther from 1 than in settings with common measures across ages (e.g., Agostinelli & Wiswall, 2016a). In this case, a 1% increase in initial skills increases early life skills by 0.2%. Assignment to IHDP treatment increases early life skills by 0.48%. Thus, an increase of 2.5% or 2.5 control-group standard deviations of initial skills would generate the same impact as random assignment to IHDP treatment—this interpretation is only a thought experiment because initial skills are a (pre-treatment) baseline variable.

There are two salient reasons for potential bias in childcare and parenting. First, standard selection bias would bias upwards the OLS elasticities. For instance, ϵ could reflect more awareness by parents (e.g., greater unobserved parenting skills, capacity, or motivation),

Table 4
Main estimates: Elasticities and predicted treatment effects at the mean.

	(1)	(2)	(3)
	Control Set 1	Control Set 2	Control Set 3
Panel a. OLS			
Elasticity (standard error)			
<i>Initial Skills</i>	0.303 (0.027)	0.289 (0.025)	0.263 (0.025)
<i>Childcare</i>	0.011 (0.001)	0.011 (0.002)	0.011 (0.001)
<i>Parenting</i>	0.458 (0.032)	0.389 (0.032)	0.309 (0.036)
Elasticity-Predicted % of Treatment Effect [<i>p</i> -value]			
<i>Initial Skills</i>	−0.297 [0.558]	−0.284 [0.557]	−0.258 [0.556]
<i>Childcare</i>	72.538 [0.000]	77.091 [0.000]	77.049 [0.000]
<i>Parenting</i>	13.175 [0.000]	11.191 [0.000]	8.899 [0.000]
Panel b. IV			
Elasticity (standard error)			
<i>Initial Skills</i>	0.207 (0.058)	0.192 (0.055)	0.195 (0.055)
<i>Childcare</i>	0.010 (0.005)	0.010 (0.005)	0.009 (0.005)
<i>Parenting</i>	0.774 (0.149)	0.752 (0.184)	0.573 (0.384)
Elasticity-Predicted % of Treatment Effect [<i>p</i> -value]			
<i>Initial Skills</i>	−0.203 [0.569]	−0.188 [0.562]	−0.192 [0.570]
<i>Childcare</i>	69.207 [0.063]	70.068 [0.054]	62.867 [0.091]
<i>Parenting</i>	22.280 [0.000]	21.652 [0.000]	16.508 [0.081]

Note: Column (1) of Panel (a) displays the elasticities (at the mean) of initial skills, childcare, and parenting and their corresponding standard error. Elasticities are based on OLS estimates of Eq. (2). They represent the change in early life skills given a 1% change in the corresponding variable. The elasticities are conditional on Control Set 1. Panel a. then displays the % of the average treatment-control difference displayed in Panel (b) of Fig. 1 explained by the elasticity and experimentally induced change in the corresponding variable. Column (1) of Panel b. is analogous in format to Column (1) of Panel a., based on IV estimates of Eq. (2). Columns (2) and (3) of Panels a. and b. are analogous in format to Column (1) for Control Sets 2 and 3. Control Set 1: initial skills and site indicators. Control Set 2: Control Set 1 and male, black, and Hispanic indicators. Control Set 3: Control Set 2 and the maternal and household characteristics in Panel b. of Table 1. Standard errors and (one-sided) *p*-values are bootstrapped, clustered at the child level, and stratified by site, treatment status, and birthweight group (low-low or high-low).

who would then mechanically invest more in their children through childcare and parenting. Second, measurement error could attenuate OLS elasticities. Childcare is driven by random assignment to treatment (control-group children spend a minor amount of time in childcare centers), and we have comprehensive measures. Consistent with this, the OLS and IV estimates are similar, indicating that selection and measurement biases are potentially minor in the case of this input. For parenting, both biases are plausible: (i) while random assignment to IHDP treatment increases this input, the control counterfactual is not minor; and (ii) while we correct parenting for measurement error, the HOME questionnaires are not infallible given the difficulty of measuring all what parents invest in their children. The elasticity estimates indicate that attenuation could be a more prominent issue, as the IV estimate of the elasticity is about two times the corresponding OLS estimate.

To further interpret the elasticities, we display the elasticity-predicted percentage of the treatment effect by input. We calculate the experimentally induced percentage change in the input, multiply it by its elasticity, and divide the resulting product by the experimentally induced percentage change in early life skills. This exercise provides

us with the percentage of the treatment effect explained by each of the inputs, as predicted by our elasticity estimates (i.e., a form of causal mediation analysis; Heckman & Pinto, 2015). The more directly impacted input, childcare, explains a greater fraction of the treatment effect on early life skills. However, the contribution of parenting is substantial, confirming parenting as an important channel through which effective early childhood education programs impact their participants. Notably, childcare and parenting explain 87% of the treatment effect on early life skills, leaving little room for a residual or omitted inputs in the production technology we postulate.

4.1. Complementarity and substitutability patterns

Fig. 2 displays the elasticities across the distribution of initial skills, holding childcare and parenting at the average. The figure indicates that childcare is more effective for relatively advantaged children. The opposite is true for parenting. Other estimates of the childcare elasticity are unavailable in the literature, as our paper is new in including this input. Available estimates of the parenting elasticity are consistent with our findings. Summarizing the literature, Heckman and

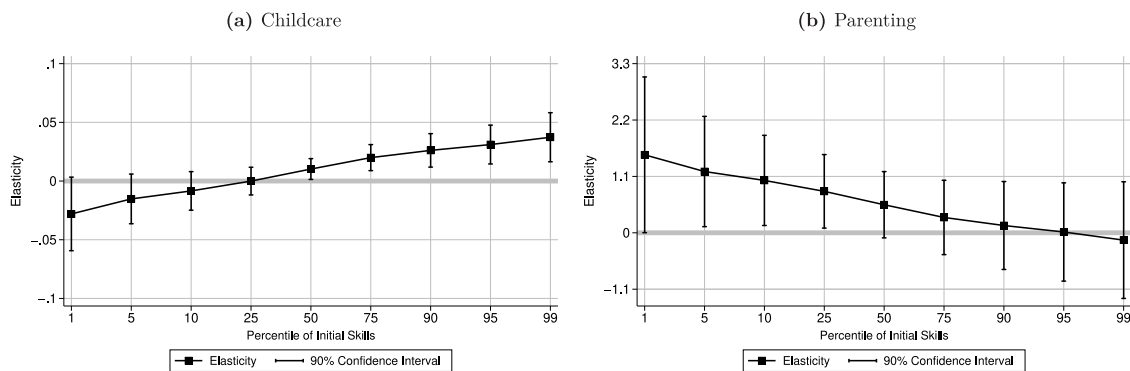


Fig. 2. Main estimates: Elasticities across the distribution of initial skills.

Note: Panel (a) displays the childcare elasticity at the percentile of initial skills in the horizontal axis, holding childcare and parenting at the mean. The elasticity represents the change in early life skills given a 1% change in childcare. It is based on IV estimates of Eq. (2), conditional on initial skills and site indicators; indicators of male, black, and Hispanic; and the maternal and household characteristics in Panel b. of Table 1. Panel (b) is analogous in format to Panel (a) for the parenting elasticity. Each elasticity is displayed with its 90% confidence interval, which is bootstrapped, clustered at the child level, and stratified by site, treatment status, and birthweight group (low-low or high-low).

Mosso (2014) state: “the empirical literature [...] is consistent with the notion that investments and endowments are direct substitutes”. In their discussion, investment refers to parenting. The authors document that the elasticity becomes positive over time, making early life a sensible period for investing in disadvantaged children.

A novelty of our paper is estimating a production technology in the context of a high-quality early childhood education program. García and Heckman (2023) argue that increasing parenting is essential to the success of these programs.¹⁶ Panel (b) of Fig. 3 shows treatment effects on this input across the distribution of skills.¹⁷ More disadvantaged children, as measured by their initial skills, enjoy the larger treatment effects. For these children, treatment assignment to IHDP sizably increases their parenting. It turns out that parenting is the more effective input for them. This finding micro-founds a recent generality documented in García and Heckman (2023): high-quality early childhood education programs boost parenting for the most disadvantaged, and this input is the more effective in the production of their early life skills, which in turn determines lifetime success.

A Counterfactual Illustration. A stylized counterfactual exercise in Panels (c) and (d) of Fig. 3 illustrates the importance of childcare and parenting in promoting mobility. We use the control group for this exercise, displaying results for the pooled sample and for the sample of disadvantaged children (below the 25th percentile of the distribution of initial skills). For disadvantaged children, the average of early life skills is 99.3 (gray dashed horizontal line), 0.7 standard deviations below the overall average of 100 (black dashed horizontal line). We then use the elasticities in Table 4 to impute percentage changes in childcare and parenting. The observed and counterfactual averages coincide when the percentage changes imputed are 0. For childcare, the elasticity is negative for the most disadvantaged. Thus, a greater imputed treatment effect leads to a lower average. For the overall sample, the average increases after imputation, which is expected given the positive impact of assignment to treatment in IHDP.¹⁸ For parenting, the elasticity

decreases as the level of initial skills increases. Imputing 1.2 of a standard deviation of parenting, equal to a 1.2% increase relative to the control-group mean, the disadvantaged children virtually catch up to the average in the overall sample under the same imputation. This percentage change happens to be what assignment to treatment generates for the most disadvantaged among the disadvantaged, which are those below the 5th percentile of the distribution of initial skills (see Panel (b) of Fig. 3).

4.2. Longer-term estimates

Following the previous discussion, elasticity estimates later in life would be informative regarding the optimal timing for investment across the distribution of initial skills. We perform an exploratory exercise given that data on inputs after age three are unavailable, and data on skills suffer from a larger fraction of non-response. We construct a latent of later childhood skills, following the same procedure outlined for early life skills and using measures of IQ at ages five and eight (we observe age-appropriate versions of the Peabody Picture and Wechsler IQ tests at each of these two ages). The sample size drops from 644 to 575. However, we can estimate a production function identical to that in Eq. (2), replacing early life skills with later childhood skills as output and with the exact same right-hand side. This estimation speaks to the productivity of the investments at age three on later childhood skills.¹⁹

While statistical tests comparing the elasticities in Fig. 2 to the elasticities from this longer-term exercise are implausible, Fig. 4 indicates a qualitative decrease in magnitude in either of the elasticities across the distribution of skills. Therefore, while we rely on the literature in arguing that earlier investments are more effective than later investments, we can conclude that (i) parenting is more effective for more disadvantaged children, as is boosted by assignment to IHDP treatment, which mimics the treatment of other high-quality early childhood education programs; and (ii) the productivity of inputs decreases over time, indicating that effective investments are to be reinforced across childhood—such reinforcement would be more effective in the presence of dynamic complementarity.

¹⁶ Parental investments are studied in the seminal works of Becker and Tomes (1979, 1986) and Behrman et al. (1982). Doepke et al. (2019) and Francesconi and Heckman (2016) are recent discussions.

¹⁷ An alternative to explore heterogeneity in the elasticities would be to provide an account of the compliers to our instruments (e.g., Abadie, 2003). Compliers are usually characterized in program-evaluation contexts with one outcome and one instrument. Our context is more complex: we consider an outcome to be a function of multiple endogenous variables and use multiple instruments. We thus explore heterogeneity by estimating elasticities across the distribution of initial skills.

¹⁸ In addition, note that this exercise holds initial skills and parenting fixed. Thus, the treatment effect on early life skills is greater than what the black solid line suggests.

¹⁹ Table A.2 in the Appendix indicates that, in terms of the observed average for the treatment and control groups, the sample of 575 individuals for this exercise is virtually identical to the main sample analyzed above. This close similarity holds for the child, maternal, and household characteristics, as well as for childcare, parenting, and the measures of early life skills that we use for constructing the early life skills latent.

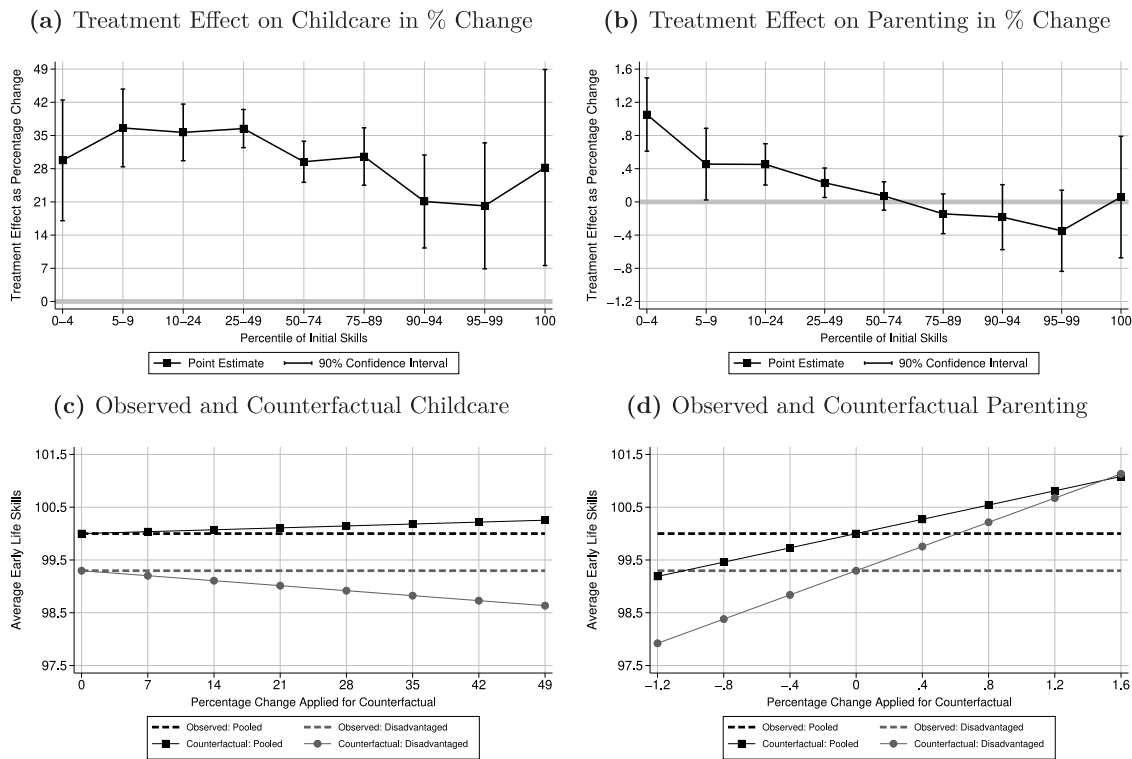


Fig. 3. Counterfactual estimates: Imputing treatment effects across the distribution of initial skills.

Note: Panel (a) displays the average treatment-control difference in childcare, as a percentage of the control-group mean, at different percentiles of initial skills. Panel (b) is analogous in format to Panel (a) for parenting. Panel (c) displays the observed means of early life skills for the pooled (dashed black) and disadvantaged (dashed gray) control-group samples. It then displays the counterfactual average predicted by estimates of Eq. (2) if the percentage change of childcare in the horizontal axis were applied to all the sample. This counterfactual is displayed for the pooled (black circles) and disadvantaged (gray circles) samples. In either case, the percentage change in childcare is applied homogeneously, but the elasticities are allowed to vary by percentile as in Fig. 2. Panel (d) is analogous in format to Panel (c) for percentage-change applications in parenting. Disadvantaged is defined as being below the 25th percentile of initial skills. Panels (c) and (d) are based on IV estimates of Eq. (2), conditional on initial skills and site indicators; indicators of male, black, and Hispanic; and the maternal and household characteristics in Panel b. of Table 1. The confidence intervals in Panels (a) and (b) are robust and clustered at the child level.

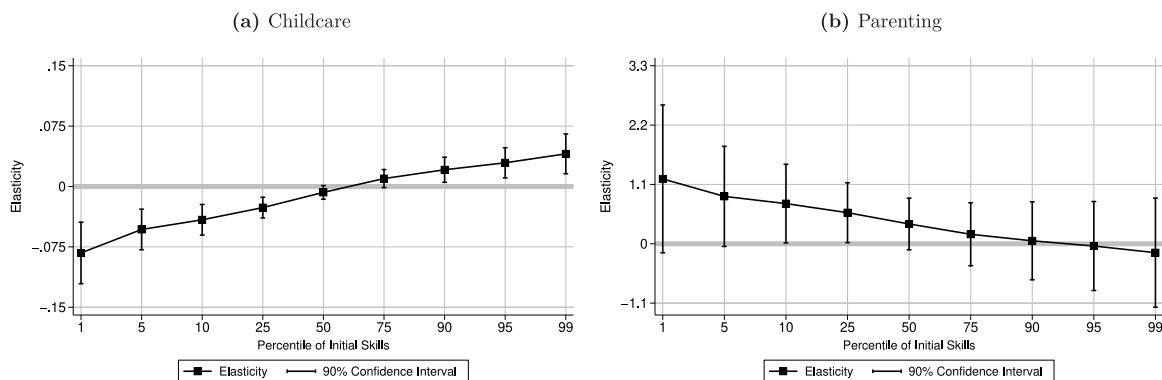


Fig. 4. Longer-term estimates: Elasticities across the distribution of initial skills.

Note: Panel (a) displays the childcare elasticity at the percentile of initial skills in the horizontal axis, holding childcare and parenting at the mean. The elasticity represents the change in later childhood skills given a 1% change in childcare. It is based on IV estimates of Eq. (2), conditional on initial skills and site indicators; indicators of male, black, and Hispanic; and the maternal and household characteristics in Panel b. of Table 1. Panel (b) is analogous in format to Panel (a) for the parenting elasticity. Each elasticity is displayed with its 90% confidence interval, which is bootstrapped, clustered at the child level, and stratified by site, treatment status, and birthweight group (low-low or high-low).

5. Final comments

A few qualifications are in order before concluding. Our results are naturally subject to limitations. The three limitations that we judge as most relevant are the following. First, the time span for our main analysis is limited, as we only estimate a short-term production technology and only observe inputs up to age three. An appealing feature of our framework is that we can base our estimates on detailed

information on inputs and outputs. A limitation, however, is that the follow-ups beyond age three have limited information on inputs (at ages five and eight) and insufficient information on inputs and skills at age eighteen. Second, our results are based on a single dimension of skills (cognition). This is common when studying skill production technologies as early in life as we do (e.g., Heckman & Zhou, 2023), although there are some exceptions (e.g., Attanasio et al., 2020a). Finally, and perhaps most importantly, the generalization of our results should

Table A.1

Summary statistics: Sample with fully observed basic baseline characteristics.

	(1)	(2)	(3)	(4)
	Control (mean)	Treatment (mean)	Difference (2) – (1)	Standard error (of difference)
Panel a. Child at Baseline				
<i>Male</i>	0.488	0.504	0.016	0.035
<i>Black</i>	0.523	0.519	–0.003	0.035
<i>Hispanic</i>	0.106	0.107	0.001	0.021
<i>Gestational Age (weeks)</i>	32.976	33.060	0.083	0.185
<i>Birth Weight (kilograms)</i>	1.784	1.822	0.038	0.032
Panel b. Mother at Baseline				
<i>Age</i>	24.766	24.454	–0.312	0.420
<i>Works</i>	0.351	0.343	–0.008	0.033
<i>Married</i>	0.481	0.415	–0.066	0.034

Note: Columns (1) and (2) display the average of the variable indicated in the row for the control and treatment groups. Column (3) displays the difference between Columns (2) and (1), and Column (4) displays the robust standard error of this difference clustered at the child level. All computations are based on the sample of 882 singleton participant children for whom we observed all the baseline characteristics listed in this table.

Table A.2

Summary statistics: Sample with outcomes observed at ages five and eight.

	(1)	(2)	(3)	(4)
	Control (mean)	Treatment (mean)	Difference (2) – (1)	Standard error (of difference)
Panel a. Child at Baseline				
<i>Male</i>	0.494	0.476	–0.018	0.043
<i>Black</i>	0.523	0.520	–0.003	0.043
<i>Hispanic</i>	0.089	0.088	–0.001	0.024
<i>Gestational Age (weeks)</i>	33.106	33.035	–0.071	0.225
<i>Birth Weight (kilograms)</i>	1.809	1.813	0.004	0.039
<i>Bayley Mental Development Index</i>	111.310	111.322	0.011	1.327
Panel b. Mother and Household at Baseline				
<i>Age</i>	24.865	24.608	–0.257	0.516
<i>IQ-Peabody Picture Vocabulary Test</i>	81.578	82.568	0.991	1.796
<i>Education</i>	12.632	12.352	–0.280	0.210
<i>Works</i>	0.388	0.396	0.009	0.042
<i>Married</i>	0.489	0.445	–0.044	0.043
<i>Other Children in Household</i>	0.701	0.692	–0.010	0.084
<i>Employed Adults in Household</i>	1.095	1.035	–0.060	0.064
Panel c. Childcare and Parenting (Ages 1 to 3)				
<i>Childcare Hours per Week</i>	5.908	27.774	21.866	1.102
<i>Parenting</i>	100.000	100.176	0.176	0.079
Panel d. Measures of Early Life Skills (Ages 2 to 3)				
<i>Bayley Mental Development Index</i>	96.716	107.026	10.311	1.553
<i>IQ-Peabody Picture Vocabulary Test</i>	85.233	94.771	9.538	1.496
<i>IQ-Stanford-Binet Test</i>	85.805	92.938	7.134	1.469

Note: Columns (1) and (2) display the average of the variable indicated in the row for the control and treatment groups. Column (3) displays the difference between Columns (2) and (1), and Column (4) displays the robust standard error of this difference clustered at the child level. All computations are based on an analysis sample of 575 singleton participant children (observed with complete item response up to age eight).

be cautious. Although IHDP included a more socioeconomically diverse sample than other early childhood education programs like the Carolina Abecedarian and Perry Preschool Projects, it is still the case that the eligibility requirements targeted the program to children with specific disadvantages. This targeting implies that the relatively disadvantaged children we discuss results for are especially disadvantaged.

Despite its limitations, our study is new in estimating a production technology in a randomized setting within the United States. Unlike other work, we include both childcare and parenting as inputs of this technology. Our results micro-found the finding that effective, high-quality early childhood education programs boost parental investment (time and resources parents spend on their children) while providing high-quality center-based programs. The boost in parenting is especially pronounced for the most disadvantaged children, for whom this input is more effective than childcare at producing early life skills. Parenting is more effective for them than for more advantaged children, making it an essential input for promoting the mobility of early life skills and, in turn, lifetime outcomes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A

See Fig. A.1 and Tables A.1 and A.2.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.econedurev.2024.102557>.

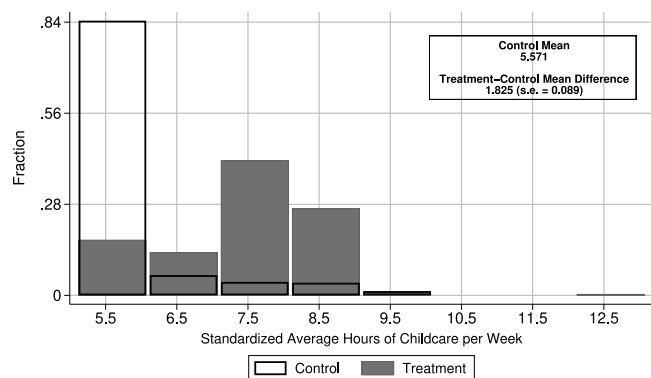


Fig. A.1. Standardized childcare (Ages 1 to 3).

Note: This figure is a version of Panel (c) of Fig. 1. In this figure, we standardize the measure of childcare by subtracting the control-group mean from it, dividing it by the control-group standard deviation, and adding the control-group mean to relocate it. We display the control-group mean, the average treatment-control difference of this standardized measure, and the robust standard error of this difference clustered at the child level.

References

- Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113, 231–263.
- Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2022). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1), 1–35.
- Agostinelli, F., & Wiswall, M. (2016a). *Estimating the technology of children's skill formation: Technical report w22442*, Cambridge, MA: National Bureau of Economic Research.
- Agostinelli, F., & Wiswall, M. (2016b). *Identification of dynamic latent factor models: The implications of re-normalization in a model of child development: Technical report w22441*, Cambridge, MA: National Bureau of Economic Research.
- Aizer, A., & Cunha, F. (2012). *The production of human capital: Endowments, investments and fertility: Technical report w18429*, Cambridge, MA: National Bureau of Economic Research.
- Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4).
- Attanasio, O., Cattani, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2020). Estimating the production function for human capital: Results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1), 48–85.
- Attanasio, O., Meghir, C., & Nix, E. (2020). Human capital development and parental investment in India. *Review of Economic Studies*, 87(6), 2511–2541.
- Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6), 1153–1189.
- Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, 4(3, Part 2), S1–S39.
- Behrman, J. R., Pollak, R. A., & Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, 90(1), 52–73.
- Bennhoff, F. H., García, J. L., & Leaf, D. E. (2024). The dynastic benefits of early-childhood education: Participant benefits and family spillovers. *Journal of Human Capital*, 18(1), 44–73.
- Bernal, R., & Keane, M. P. (2011). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics*, 29(3), 459–512.
- Boardman, J. D., Powers, D. A., Padilla, Y. C., & Hummer, R. A. (2002). Low birth-weight, social factors, and developmental outcomes among children in the United States. *Demography*, 39, 353–368.
- Boulet, S. L., Schieve, L. A., & Boyle, C. A. (2011). Birth-weight and health and developmental outcomes in US children, 1997–2005. *Maternal and Child Health Journal*, 15, 836–844.
- Bradley, R. H., & Caldwell, B. M. (1984). The HOME inventory and family demographics. *Developmental Psychology*, 20(2), 315.
- Bradley, R. H., Caldwell, B. M., Brisby, J., Magee, M., Whiteside, L., & Rock, S. L. (1992). The HOME inventory: A new scale for families of pre and early adolescent children with disabilities. *Research in Developmental Disabilities*, 13(4), 313–333.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37(1), 62–83.
- Chaparro, J., Sojourner, A., & Wiswall, M. (2020). *Early childhood care and cognitive development: Technical report w26813*, Cambridge, MA: National Bureau of Economic Research.
- Chen, J., & Roth, J. (2024). Logs with zeros? Some problems and solutions. *Quarterly Journal of Economics*.
- Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4), 738–782.
- Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In *Handbook of the economics of education: Vol. 1*, (pp. 697–812). Elsevier.
- Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883–931.
- Del Bono, E., Kinsler, J., & Pavan, R. (2022). Identification of dynamic latent factor models of skill formation with translog production. *Journal of Applied Econometrics*, 37(6), 1256–1265.
- Doepke, M., Sorrenti, G., & Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, 11, 55–84.
- Donald, S. G., & Newey, W. K. (2001). Choosing the number of instruments. *Econometrica*, 69(5), 1161–1191.
- Dougan, W. R., García, J. L., & Polovnikov, I. (2023). *High-quality early-childhood education at scale: Evidence from a multisite randomized trial: Technical report w31694*, Cambridge, MA: National Bureau of Economic Research.
- Elango, S., García, J. L., Heckman, J. J., & Hojman, A. (2016). In R. A. Moffitt (Ed.), *Economics of means-tested transfer programs in the United States: Vol. 2, Early childhood education* (pp. 235–297). Chicago: University of Chicago Press, Chapter 4.
- Francesconi, M., & Heckman, J. J. (2016). Child development and parental investment: Introduction. *The Economic Journal*, 126(596), F1–F27.
- Freyberger, J. (2022). *Normalizations and misspecification in skill formation models: Technical report*, arXiv, arXiv:2104.00473.
- García, J. L., & Heckman, J. J. (2023). Parenting promotes social mobility within and across generations. *Annual Review of Economics*, 15, 349–388.
- García, J. L., Heckman, J. J., Leaf, D. E., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood education program. *Journal of Political Economy*, 128(7), 2502–2541.
- Gross, R. T., Spiker, D., & Haynes, C. W. (1997). *Helping low birth weight, premature babies: The infant health and development program*. Stanford University Press.
- Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6(1), 689–733.
- Heckman, J. J., & Pinto, R. (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric Reviews*, 34(1–2), 6–31.
- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6), 2052–2086.
- Heckman, J., & Zhou, J. (2023). *Nonparametric tests of dynamic complementarity: Technical report*, Center for the Economics of Human Development, Working Paper.
- Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1), 97–126.
- National Bureau of Economic Research (2022). Vital statistics natality birth data. URL <https://www.nber.org/research/data/vital-statistics-natality-birth-data>.
- National Center for Health Statistics (2022). National vital statistics system. URL <https://www.cdc.gov/nchs/nvss/births.htm>.
- Newey, W. K. (1990). Efficient instrumental variables estimation of nonlinear models. *Econometrica*, 58(4), 809, arXiv:2938351.
- Pavan, R. (2016). On the production of skills and the birth-order effect. *Journal of Human Resources*, 51(3), 699–726.
- Ramey, C. T., Bryant, D. M., Wasik, B. H., Sparling, J. J., Fendt, K. H., & La Vange, L. M. (1992). Infant health and development program for low birth weight, premature infants: Program elements, family participation, and child intelligence. *Pediatrics*, 89(3), 454–465.
- Rosenzweig, M. R., & Wolpin, K. I. (1988). Heterogeneity, intrafamily distribution, and child health. *Journal of Human Resources*, 23(4), 437, arXiv:145808.
- Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. In *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*: In *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*: Vol. 6.109–120.
- Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from head start. *American Economic Journal: Applied Economics*, 7(4), 76–102.
- Wang, J., & Sauer, M. V. (2006). In-vitro fertilization (IVF): A review of three decades of clinical innovation and technological advancement. *Therapeutics and Clinical Risk Management*, 2(4), 355.
- Williams, B. (2020). Identification of the linear factor model. *Econometric Reviews*, 39(1), 92–109.