The Intergenerational Elasticity of Earnings: Exploring the Mechanisms

DIAL Working Paper Series 11/2021

Uta Bolt, Eric French, Jamie Hentall MacCuish, Cormac O'Dea

Dynamics of Inequality Across the Lifecourse: structures and process

dynamicsofinequality.org

NORFACE NETWORK

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 724363

The Intergenerational Elasticity of Earnings: Exploring the Mechanisms *

Uta Bolt,† Eric French‡‡ Jamie Hentall MacCuish§ and Cormac O'Dea¶

August 25, 2021

Abstract

Using data covering a single cohort's first 55 years of life, we show that most of the intergenerational elasticity of earnings (IGE) is explained by differences in: years of schooling, cognitive skills, investments of parental time and school quality, and family circumstances during childhood. To decompose the fraction of the IGE explained by each of these channels, we implement a multi-level mediation analysis combined with a latent factor framework that accounts for measurement error. Multilevel mediation analysis allows us to assess not only the *direct* effect of each channel on the IGE, but also its *indirect* effects working through the other channels, thus providing an in-depth understanding of the link between parents' and children's earnings. Of these channels, we show that the main driver of the IGE is increased levels of parental investments received by children of high income parents early in their lives, which encourages greater cognitive development and lifetime earnings.

Keywords: Parental Investments, Cognitive Skills, Intergenerational Elasticity of Earnings

JEL codes: I24, J24, C38

*We are very grateful to Jaime Arellano-Bover, Jo Blanden, Paul Gregg, Lindsey Macmillan, Vitor Possebom, Lindsey Uniat and the CSaP Horn Fellows roundtable for helpful discussions and Sreevidya Ayyar for excellent research assistance. Funding from the Social Security Administration through the Michigan Retirement Research Center (MRRC grant UM17-15) and the Economic and Social Research Council ("Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies" (RES-544-28-50001), "Inequality and the insurance value of transfers across the life cycle" (ES/P001831/1)), "Intergenerational transfers, insurance, and the transmission of inequality" (ES/V001248/1), and NORFACE for Grant 462-16-120: Trends in Inequality: Sources and Policy" for this work is gratefully acknowledged. We are grateful to the Centre for Longitudinal Studies (CLS) for the use of these data and to the UK Data Service for making them available. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA or ESRC. Any errors are the authors' own.

†University College London and Institute for Fiscal Studies

‡University of Cambridge, University College London, and Institute for Fiscal Studies

§University College London and Institute for Fiscal Studies

¶Yale University, NBER and Institute for Fiscal Studies

1 Introduction

Income across generations is persistent. A developing literature investigates the mechanisms behind this persistence, such as differences in parental time, years of schooling, and family background.[1] However, there is still scarce evidence on the quantitative importance of each of these mechanisms for the intergenerational elasticity of earnings (IGE) and on the way they interact with each other. In this paper we provide such evidence.

Using a unique longitudinal data set covering a British cohort's first 55 years of life, we show that 54% of the IGE for males, and 62% for females, can be accounted for by differences in years of schooling, cognitive skills, parental investments in time and school quality, and family background. Of these channels, we show that parental investments are the main driver of the link between parents' income and children's earnings via higher cognitive development.

Our approach is to decompose the IGE at different levels of mediation, allowing us to quantify not only the impact of different factors (e.g., parental time and school quality investments) on lifetime earnings, but also to understand the mechanisms by which these factors impact earnings (e.g., through cognitive skills and years of education).[2] Specifically, we estimate how parental income affects each of the factors that we consider, and how these, directly or indirectly through the other factors, affect an individual's lifetime earnings. Similar to Heckman et al. (2018), we do not explicitly model agents' preferences, but instead we approximate their decision rules, and exploit restrictions arising from the lifecycle timing of human capital development.

We allow for four distinct mechanisms that can generate persistence in earnings across generations. The first mechanism is years of schooling. We find that the higher levels of schooling of children from richer families matter for the IGE.

However, once we consider the second mechanism, cognitive skills, we find that the effect of years of schooling on the IGE can be entirely explained by differences in age 16 cognition. Cognition, therefore, is crucial - it directly affects earnings, and it is the main driver of schooling decisions.

The third mechanism is investments. Richer families invest more time in their children and also send their children to better quality schools. These investments explain 47% (46%) of the IGE for males (females) once we allow investments to subsequently affect the development of cognition and years of schooling. Both types of investment are important, with time investments playing a more important role

[1] See Solon (1992), Dearden et al. (1997), Mazumder (2005), Chetty et al. (2014) for evidence on the intergenerational correlation of earnings. The importance of time investments in children for their later life earnings has been extensively studied (Cunha and Heckman (2008) , Daruich (2018), Lee and Seshadri (2019)) and it has been well documented that richer parents invest more time in their children (Guryan et al. (2008)). Aizer and Cunha (2012) and Del Boca et al. (2014) show that richer parents have fewer children, allowing them to invest more time in each child.

[2] See Klein and Goldberger (1955) and Theil (1958) for early applications of mediation analysis in economics, and Blanden et al. (2007), Heckman et al. (2013), Conti et al. (2016), Mogstad et al. (2021) for recent relevant applications.

early in childhood, and school quality investments becoming more important later in childhood. Investments are critical for understanding later life earnings, but their importance arises through their effect on cognitive skills and years of schooling. We find little evidence that these investments are important for earnings over and above their impact via cognitive skills. This is an interesting finding in its own right, as it is typically *assumed* that investments impact later life outcomes only through their impact on observable skills. However, this assumption is rarely tested (see Heckman et al. (2013) for an example that does test this assumption). We fail to reject this key assumption.

The fourth is family background. Parental education and family size, through their impact on investments, explain 19% (34%) of the IGE for males (females). However, we also find that differences in investments between high and low income parents are not entirely driven by family background differences. The direct link between parental income and investments explains 33% (25%) of the IGE. These results are consistent with the view that if financial constraints are important for understanding the persistence in income across generations, it is because they constrain early-in-life investments, not years of schooling.

Whilst other papers have considered the channels above individually, or jointly by matching moments from different data sources, to the best of our knowledge, this is the first paper to evaluate these major drivers of the IGE using a single sample of people. Using the same sample throughout our mediation analysis enables us to measure inputs and outputs of one group of people in a single setting, and also to test whether early life investments (e.g. parental time) have independent effects on late life outcomes, over and above their effects on intermediate outcomes (e.g. cognition). Our mediation analysis also builds on recent advances in latent factor methods and carefully takes into account measurement error. This turns out to be important - ignoring measurement error attenuates the fraction of the IGE explained by parental investment by 45%.

The rest of this paper proceeds as follows. Section 2 relates our work to previous literature. Section 3 describes the data and documents descriptive statistics on schooling, cognition, investments, and family background. Section 4 outlines our mediation approach and Section 5 provides results and robustness checks. Section 6 concludes.

2 Literature

This paper focusses on understanding the importance of candidate mechanisms for the intergenerational elasticity in earnings. The setting for our empirical work is the UK, where the level of intergenerational persistence is similar to that observed in the US, and substantially higher than in most developed countries

(Corak (2013)).[3]

A growing literature examines what drives these intergenerational associations (see Black et al. (2011)). Here, we summarize the literature on the intergenerational elasticity of earnings and its interplay with the mechanisms we explore: years of schooling, cognition, parental time investments, school quality and family background.

Consider the role of schooling first. Children of richer parents spend on average more years in education which leads to higher earnings. Two distinct channels have been emphasized in interpreting this. One emphasizes that children of richer parents develop greater cognitive skills, which keeps them in school longer (Keane and Wolpin (1997), Carneiro and Heckman (2002)). Another emphasizes the role of borrowing constraints (Lochner and Monge-Naranjo (2012), Lee and Seshadri (2019), Caucutt and Lochner (2020)). Our mediation analysis allows us to investigate the extent to which the effect of schooling on the IGE is mediated by cognition, and to what extent it can be directly accounted for by parental income.

The role of differences in the skills between children of low and high income parents has also been previously highlighted. Exploiting plausible variation in income, a number of papers (e.g Milligan and Stabile (2011), Dahl and Lochner (2012) and Agostinelli and Sorrenti (2018)) find that increases in parental income have a positive effect on children's test scores. Deckers et al. (2017) also show that children of high income parents have higher cognition.[4] Our framework allows us to quantify how much of the differences in cognition can be explained by differential investments or differences in parental education, and to what extent they can be explained directly by parental income.

Papers in a large recent literature (e.g. Cunha and Heckman (2008), Cunha et al. (2010), Attanasio et al. (2020)) estimate production functions of human capital and have shown the formative role of parental time investments in the development of human capital. This evidence, combined with the fact that high income people spend more time with their children (Guryan et al. (2008)), means that parental differences in time investments are likely to be important in propagating correlations in earnings across generations. We estimate how parental investments affect cognitive skills at age 16 using latent factor methods and incorporate this into our mediation analysis. We find that differences in parental time investments are the dominant channel driving the IGE, at least for men. Applications of the production functions estimated in this literature often assume that investments affect earnings *only* through their impact on human capital. We can test whether this assumption holds, or whether parental investments can affect children's lifetime earnings above and beyond their effects on cognition. We fail to reject this assumption.

The literature on the role of school quality is smaller than that on time investments. Altonji and

[3] Several papers estimate this elasticity using the same NCDS data we use, and find it to be between 0.21 and 0.43, depending on the measures and procedures used (Dearden et al. (1997), Gregg et al. (2016), Belfield et al. (2017))

[4] Non-cognitive skills have also been widely studied (Papageorge et al. (2019), Heckman et al. (2006)), though their effects on earnings are usually estimated to be smaller than for cognition.

Dunn (1996) and Dearden et al. (2002) find modest effects of school quality on children's outcomes. More recently, Chetty et al. (2014) show that where a household lives has important impacts on income mobility, which may be related to school quality. Our data contains several measures of school quality, and allows us to look separately at the role of both parental time and school quality investments. We find that differences in school quality investments are the dominant channel driving the IGE, at least for women.

Turning finally to the role of family background, studies have investigated the impact of growing up in larger families (Black et al. (2005), Angrist et al. (2010), Bhalotra and Clarke (2020)), as well as effects of having more educated parents (Meghir and Palme (2005), Nybom and Stuhler (2014)). Whilst these papers are able to estimate the causal effect of family background on child outcomes by using plausibly exogenous variation, our aim is to trace mechanisms through which any family background effect operates. In particular, we can estimate whether these family characteristics affect lifetime earnings through improved cognition and schooling or through parental investments.

The two papers most similar to ours are Gayle et al. (2018) and Blanden et al. (2007). Gayle et al. (2018) estimate a rich intergenerational lifecycle model of parental decision-making that includes several of the channels that we consider. Our contribution relative to that paper, and other recent work which models parental decisions over certain investments[5], is that we can use our rich data set to quantify and test how different mechanisms affect the IGE, with very limited pre-imposition as to how these channels operate. Blanden et al. (2007) estimate the extent to which cognitive skills, non-cognitive skills, educational attainment, and labor force attachment have accounted for the rise of intergenerational earnings persistence between the NCDS 1958 cohort, and a cohort born in 1970. We build on their work in three ways. First, instead of taking cognition or schooling as given, we estimate the extent to which they are affected directly by parental income, but also indirectly through differences in investments and family background. Second, we use latent factor methods as recently developed in Heckman et al. (2006), Heckman et al. (2013), and Heckman et al. (2018), in order to account for the fact that ability, investment and school quality are measured with error. Finally, we study the lifetime earnings of both males and females rather than using male income at a particular point in time as an outcome.

To summarize, the aim of this paper is to quantify the effect of various channels on the IGE with minimal restrictions beyond those imposed by the lifecycle timing of human capital development. That is, within a framework informed by that lifecycle, we use the data to speak to the importance of each channel's direct and indirect effect on the IGE. Apart from being of interest in their own right, the results of this paper can be valuable for papers attempting to explicitly model parental behaviour, by being

[5]For example Gayle et al. (2015), Lee and Seshadri (2019), Caucutt and Lochner (2020) and Daruich and Kozlowski (2020).

informative about what channels quantitatively matter, as well as how these should be included in a model.

3 Data

Our data comes from the UK's National Child Development Study (NCDS).[6] The dataset covers the full population of children born in one particular week of March 1958 and continues to follow them to this day. The NCDS is a globally-unrivalled resource for social scientists in its combination of information about investments in early childhood with information on ability, educational outcomes and later-life earnings. To the best of our knowledge, it is the only dataset in the world that provides detailed information on early-life investments and earnings outcomes over the entire working life. The data include multiple measures of children's ability, parental time investments and school quality in each wave. This allows us to posit the existence of underlying latent factors and to use a latent factor model to extract signals from noisy measures rather than assuming that cognition, investments and school quality can be adequately captured by one particular measure.

The initial survey at birth has been followed by subsequent surveys at ages: 7, 11, 16, 23, 33, 42, 46, 50 and 55.[7] The data from childhood includes information on several measures of cognition and parental investments, number of siblings, parental education, and parental income. Later waves of the study record educational outcomes, demographic characteristics, earnings, and hours of work. Table 2 gives a list of all the different cognition, time investment, and school quality investment measures we use. Details about the sample we use are given in Appendix A.2.

In the rest of this section, we first describe how we construct the measures of parental income and child earnings that we use. We then document inequalities in family background, investments, and outcomes by parental income over the lifecycle in the NCDS data.

3.1 Parental Income

When NCDS cohort members were aged 16, comprehensive data on parental income was collected. Our measure of parental income sums across: father's earnings, mother's earnings, and other income, all net of taxes. Further details can be found in Appendix A.1.1. In the descriptive table which follows, we split households by parental income tertile. Average annual income (at the time the child is 16 years old) is reported in the first row of Table 1 and ranges from £13,600 (at the 2014 price level) in the first tertile to £29,300 in the top tertile.

[6]The NCDS is provided by the Centre for Longitudinal Studies (2017) at the Institute of Education, University College London.

[7]The age-46 survey is not used in any of the subsequent analysis as it was a more limited telephone interview only.

3.2 Children's Earnings

We observe cohort members' gross earnings at ages 23, 33, 42, 50, and 55. For further details on these earnings measures, see Appendix A.1.2. Average annual earnings (setting earnings of non-workers to 0) over all ages 23-55 are reported in the second row of Table 1. The positive gradient with respect to parents' income is immediately apparent – average gross annual earnings of NCDS cohort members rises from £17,300 for those with parents in the bottom income tertile, to £20,400 for those whose parents are in the top income tertile.

Table 1: Descriptive Statistics and Means by Parental Income Tertile

			Parental Income Tertile			
Variable	Mean	SD	Bottom	Middle	Top	P-values
Income variables						
Parents' household income at 16	21,134	7,523	13,598	20,496	29,308	0.00
Children's average annual earnings	18,899	16,913	17,293	19,019	20,386	0.00
Education						
Age left school	18.0	1.8	17.9	17.9	18.1	0.02
Cognition						
Reading at age 16	0.00	1.00	-0.11	0.01	0.10	0.00
Math at age 16	0.00	1.00	-0.08	-0.02	0.10	0.00
Teacher assessed English ability at age 16	0.00	1.00	-0.08	-0.04	0.12	0.00
Teacher assessed Math ability at age 16	0.00	1.00	-0.09	-0.01	0.10	0.00
Time investment						
% of fathers very interested at age 7	21.5	41.1	18.6	21.5	24.4	0.01
% of mothers very interested at age 7	34.4	47.5	31.5	34.1	37.4	0.03
% of fathers very interested at age 11	23.5	42.4	20.1	23.8	26.4	0.01
% of mothers very Interested at age 11	33.4	47.2	29.8	34.3	36.1	0.02
% of fathers very interested at age 16	29.9	45.8	28.2	28.6	32.9	0.08
% of mothers very interested at age 16	33.3	47.1	31.5	32.8	35.6	0.19
School quality						
% whose PTA holds meetings at age 7	57.7	49.4	56.8	57.6	58.7	0.71
Student-teacher ratio at age 11	24.8	9.35	25.4	24.7	24.3	0.06
% from child's class suitable for GCEs at age 11	23.6	14.0	23.4	23.1	24.3	0.20
Student-teacher ratio at age 16	17.2	1.93	17.2	17.3	17.1	0.05
% from child's class studying for GCEs at age 16	46.3	31.2	44.0	44.4	50.5	0.00
% from child's school who complete school	55.7	25.8	53.5	54.2	59.6	0.00
Family Background						
Number of siblings	2.04	1.46	2.13	1.93	2.05	0.01
Father's age left school	15.0	1.17	14.9	14.8	15.2	0.00
Mother's age left school	15.1	1.06	15.0	15.1	15.3	0.00

Note: This table presents descriptive statistics and means by parental income tertile for parental income, child earnings, and several measures of the mediators in our analysis (years of schooling, cognition, time investments, school quality, family background). The relevant sample consists of 2,637 individuals. Income and earnings variables are reported as annual values, deflated to 2014 prices. The final column reports *p-values* from the *F-tests* testing the null hypothesis of equality of means across income tertiles.

3.3 Years of Schooling

The age at which an individual left school is constructed using the highest educational qualification recorded by the time cohort members reach age 33. The second panel of Table 1 shows that children with parents in the bottom income tertile spend 0.21 years less in school than children with parents in the highest income tertile.

3.4 Cognition

As part of the survey, cohort members took part in tests to measure their cognitive skills.[8] The third panel of Table 1 summarises standardized math and reading test scores at age 16 by parental income tertile. As one might expect, children from richer households develop greater cognitive skills; at age 16, reading scores were 21% of a standard deviation higher on average for children with parents in the highest income tertile compared to children with parents in the lowest tertile. For math scores, the difference between children from high versus low income families is 18% of a standard deviation.

Similar patterns can be found for teacher-rated English language and math ability at age 16.

3.5 Parental Investments

The NCDS has detailed measures of the parental investments which cohort members received during childhood. The full set of parental time and school quality measures as well as cognitive skill measures are listed in Table 2. These measures come from different sources – some are from surveys of parents, others from surveys of school teachers and headteachers. The fourth and fifth panels of Table 1 document, and the next subsections discuss, the income gradients for a subset of these measures; these gradients are similar across all measures (see Table 8 in Appendix A for descriptive statistics on the full set of measures).

3.5.1 Parental Time Investments

Table 1 shows teachers' responses to a question about how interested parents are in their children's education (very interested, a little interested, not interested at all), asked when the children are 7, 11 and 16. We report the fraction of mothers and fathers who are reported to be very interested. The differences by parental income are evident at all ages for both parents. For example, 32% of mothers in the bottom tertile are adjudicated to be 'very interested' in their 7-year-old's education, compared to 37% of mothers at the top. The relevant figures for fathers are 19% and 24% respectively. Table 8 in the Appendix shows that a qualitatively similar gradient between parental income and time investments is evident for our other measures of time investment.

[8]We use the terms cognitive skills and cognition interchangeably.

Table 2: List of ability and investment measures

Cognition	
Age 16	Reading score, Math score, Teacher-assessed Math and English ability
Time Investments	
Age 0-7	Teacher's assessment of parents' interest in child's education, Parent outings with child, Whether parents read to the child
Age 7-11	Teacher's assessment of parents' interest in child's education, Parent outings with child, Parents' ambitions regarding child's educational attainment
Age 11-16	Teacher's assessment of parents' interest in child's education, Parents' ambitions regarding child's educational attainment
School Quality Investments	
Age 0-7	Whether school has a parent-teacher association (PTA), whether PTA conducts educational meetings, Social class of fathers in the child's class, Type of school
Age 7-11	Class size, Student-teacher ratio, Type of school, Proportion of children in class suitable for GCEs
Age 11-16	Student-teacher ratio, Type of school, Proportion of children in child's school who: complete school / move on to full-time degrees / pass 2+ A-levels / are studying for GCEs

Note: Data collected at ages 7,11 and 16 for children includes measures of ability, non-cognitive skills, parental time investments, family circumstances, and parental income. Note, all investment measures are retrospective, so age 0-7 investments are measured at age 7, age 7-11 investments are measured at age 11, age 11-16 investments are measured at age 16.

3.5.2 School Quality

The fifth panel of Table 1 presents some measures of school quality at different ages. Children of high income parents are more likely to go to schools where: parents attend educational meetings at age 7, student-teacher ratios are low at age 11, and a high fraction of students are doing GCEs at age 16 (an optional exam for progressing to further education) and complete secondary education. However, some measures of school quality, such as the student-teacher ratio at age 16, hardly differ by parental income tertile. This highlights the advantage of our factor analytic approach, which does not focus on any individual measure of school quality.

3.6 Family Background

In our main analysis we use three family background variables: the number of siblings and mothers' and fathers' years of education (though in Section 5.5 we will show the robustness of our results to including several more). The sixth and final panel in Table 1 shows the well-known fact that poorer families tend to be larger. NCDS cohort members from households in the bottom income tertile grew up with more siblings on average than those in the other income tertiles. Mothers' education also differs significantly between parental income tertiles - mothers in the highest tertile spent on average four more months in school than those in the lowest tertile.

The next section shows how we use this rich data set to understand how the parental income gradients described here map into the gradient of lifetime earnings.

4 Approach

The goal of this paper is to decompose the IGE into fractions explained by schooling, cognition, investments, and family background using mediation analysis. The structure of this mediation framework follows logically from the lifecycle timing of human capital development. We allow factors that are determined later in life (such as years of schooling and cognition) to be affected by factors determined earlier in life (such as investments and family background).

Throughout this analysis, we correct for potential biases arising from measurement error in parental income, cognition, time investments, and school quality and use data on multiple measures of each of these last three factors. Our approach builds on recent work in economics that uses mediation analysis such as Heckman et al. (2013) and Heckman et al. (2018), and adds several more layers of mediation.[9]

The first part of this section describes our approach for estimating the IGE. The second part describes the multi-level mediation approach in more detail and illustrates the estimation of the equations for lifetime earnings, years of schooling, cognitive skill, and investments. The final part describes our latent factor approach and how we account for measurement error.

4.1 Measurement of the Intergenerational Elasticity of Earnings

To estimate the IGE, we would like to regress (de-meaned) log lifetime earnings of the NCDS cohort member $\ln Y_i$ on (de-meaned) log lifetime income of their parents $\ln Y_{Parent,i}$:

$$\ln Y_i = \rho \ln Y_{Parent,i} + u_i \tag{1}$$

where ρ represents the intergenerational elasticity of earnings. Since we consider de-meaned variables here, we omit the intercept.

We do not have a measure of parents' lifetime income and so cannot estimate ρ directly. We do, however, observe parents' income when the cohort member was 16. Thus, we estimate a regression of (de-meaned) log lifetime earnings of the NCDS cohort member $\ln Y_i$ on (de-meaned) log income of their parents measured when NCDS cohort members are 16 years old, $\ln Y_{Parent,i,16}$:

[9]For applications of mediation in economics see Blanden et al. (2007), Conti et al. (2016), Gelbach (2016), Bandiera et al. (2020) and Mogstad et al. (2021). Mediation analysis is also very common in the psychology literature, see MacKinnon et al. (2007) for an overview.

$$\ln Y_i = \rho^* \ln Y_{Parent,i,16} + u_i. \tag{2}$$

Previous research has shown that using parents' income at a point in time as a proxy for parents' lifetime income can lead to significant attenuation bias. For example, Haider and Solon (2006) find that using point in time income can bias the estimate of the IGE downwards by between 35-80% if parents' income is measured when the parents are very young or very old, although they also show that the bias is modest when the parents are in their 40s. Fortunately for our analysis, parents are on average age 43 when their income is measured.

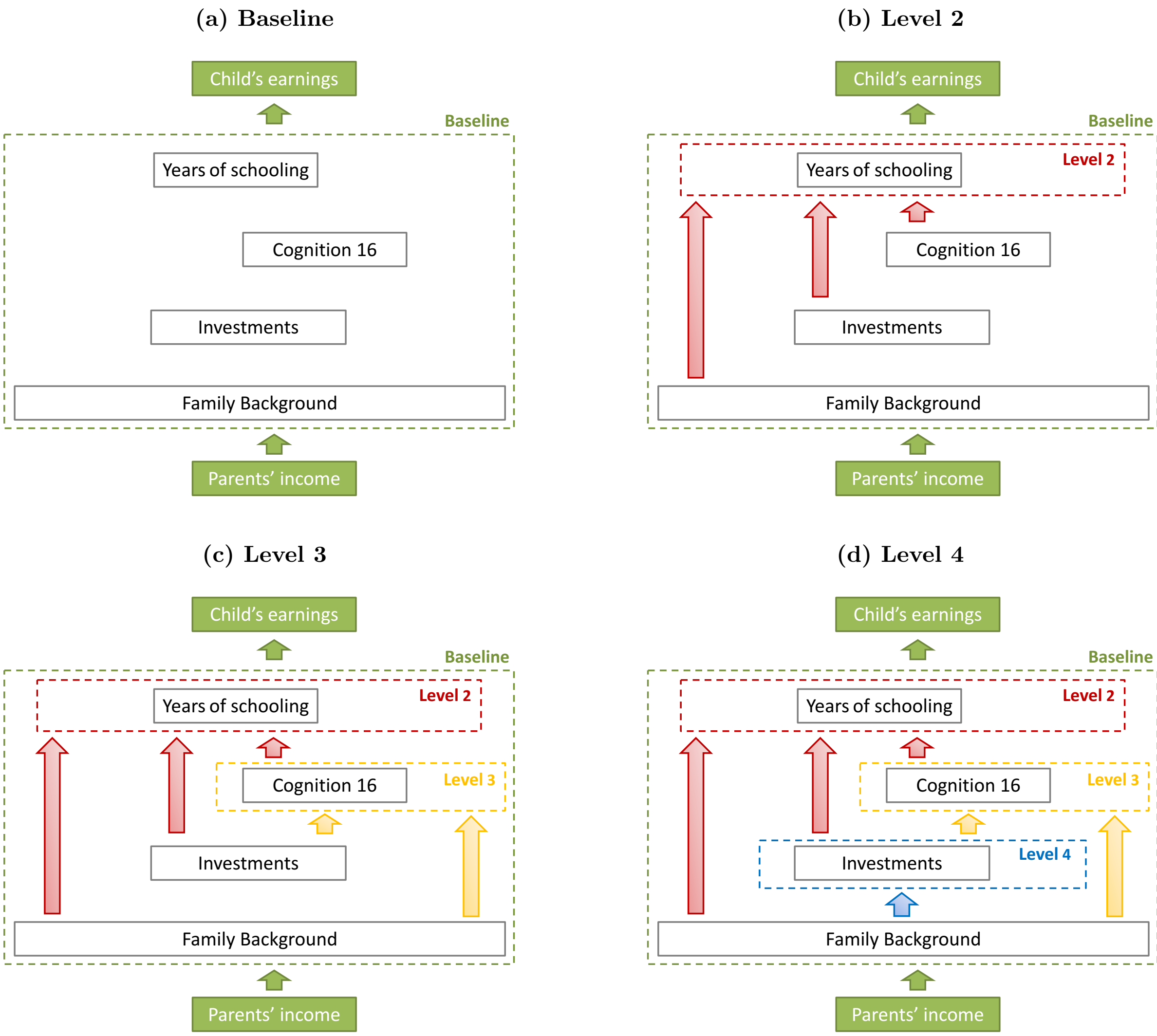
Nevertheless, we also address the issue of measurement error explicitly. We use an errors-in-variables framework that is similar to the framework adopted in Haider and Solon (2006). We estimate the attenuation bias using information on children's lifetime and point in time income at age 42. As parents are on average age 43 when we observe their income, we assume that the reliability ratio is the same across generations and use it to correct the IGE. See Appendix B for further details.

4.2 Decomposition and Multilevel Mediation Analysis

Figure 1 illustrates our decomposition and mediation approach. We begin with a baseline level (illustrated in panel a) in which we estimate the impact of parental income on the mediating factors: years of schooling, cognition, investments and family background (which are depicted in the box in panel a). Next, we estimate the impact of each of these mediating factors on the cohort member's lifetime earnings. We then decompose the IGE into the fractions explained by the different mediating factors. In this decomposition, each element explains a fraction of the IGE, holding constant the effect of all other mediators. The decomposition in level 2 (panel b) allows for family background, investments and cognition to affect the IGE directly, but also indirectly via an effect on years of schooling. The decomposition in level 3 (panel c) allows for family background and investments to affect the IGE directly and indirectly through both cognition and schooling. Finally, the decomposition in level 4 (panel d) allows for family background to affect the IGE directly and through investments, cognition, and years of schooling. The decomposition procedure extends the one in Gelbach (2016) to nest multiple layers of mediation and also to account for measurement error. An attractive feature of the decomposition is that the shares explained in the decomposition are invariant to the level of mediation. We now give more detail about each of these stages in turn.

Baseline - Direct effects on lifetime earnings To implement this multi-level decomposition, we start by estimating the parameters of the following model:

Figure 1: Overview of Mediation Approach



Note: In all levels, parents' income directly impacts all factors, which in turn affect the child's earnings. Additionally, in levels 2, 3, and 4, factors within the box impact each other as shown by the arrows.

$$\ln Y_i \quad = \quad \alpha_S S_i + \alpha_C C_i + \boldsymbol{\alpha}_I \boldsymbol{I}_i + \boldsymbol{\alpha}_F \boldsymbol{F}_i + \alpha_{Y_P} \ln Y_{Parent,i} + u_i^Y \tag{3}$$

where S is individual i's years of schooling and C is the child's cognition measured at age 16. $\boldsymbol{I} = [inv_7, inv_{11}, inv_{16}, sq_7, sq_{11}, sq_{16}]$ is a vector of all investments (time investments and school quality) at ages 7, 11, and 16. $\boldsymbol{F} = [ed_m, ed_f, sib]$ is a vector containing the family background variables (i.e. mother's and father's years of schooling, and the number of siblings of child i).

Our goal is to estimate the shares of the IGE explained by different channels. To do this, we first totally differentiate equation (3) with respect to $\ln Y_{Parent,i}$, which yields:

$$\frac{d\ln Y_i}{d\ln Y_{Parent,i}} = \alpha_S \frac{dS_i}{d\ln Y_{Parent,i}} + \alpha_C \frac{dC_i}{d\ln Y_{Parent,i}} + \boldsymbol{\alpha}_I \frac{d\boldsymbol{I}_i}{d\ln Y_{Parent,i}} + \boldsymbol{\alpha}_F \frac{d\boldsymbol{F}_i}{d\ln Y_{Parent,i}} + \alpha_{Y_P} \quad (4)$$

We then estimate how each of the covariates is individually related to parental income. For example, in the case of cognition, to calculate $\frac{dC}{d\ln Y_{Parent}}$ we estimate:

$$C_i = \kappa_C \ln Y_{Parent,i} + v_i^C \quad (5)$$

Using equations (4) and (5) we compute the total impact of parental income on earnings via cognition to be: $\frac{\partial \ln Y}{\partial C} \frac{dC}{d\ln Y_{Parent}} = \alpha_C \cdot \kappa_C$. Further, from equation (1) we know that the IGE is $\frac{d\ln Y}{d\ln Y_{Parent}} = \rho$. With these two quantities we can estimate the share of the IGE explained by cognition to be $\frac{\alpha_C \cdot \kappa_C}{\rho}$. We perform similar decompositions for all of the other variables – schooling, investments and family background.

As equation (4) shows, α_{Y_P} is the part of the IGE not explained by any of the other covariates, and thus $\frac{\alpha_{Y_P}}{\rho}$ is the total fraction of the IGE not explained by them.

Estimation of equation (3) allows us to test some restrictions in the lifetime earnings equation that are frequently made in the literature. Parental investments might impact future earnings only through their impact on cognition and subsequent school attainment. If this is the case, then they should have zero predictive power for child lifetime earnings, conditional on cognition and schooling. That is, in the language of psychology, there would be *complete mediation* (MacKinnon et al. (2007)). In fact, this is the standard assumption in the literature of dynamic skill investment (Caucutt and Lochner (2020), Lee and Seshadri (2019), Daruich (2018)). However, it is possible that, for example, school quality may have an effect on lifetime earnings above and beyond years of schooling, for instance by allowing individuals to have access to a more opportune social network and hence get better jobs. Whilst the coefficients on the covariates tell us individually whether this is the case, we also test whether groups of covariates jointly have a significant effect on lifetime earnings using *F-tests*. That is, we test for the joint significance of time investments at all ages, $\alpha_{inv_7} = \alpha_{inv_{11}} = \alpha_{inv_{16}} = 0$, school quality at all ages, $\alpha_{SQ_7} = \alpha_{SQ_{11}} = \alpha_{SQ_{16}} = 0$, and all family background variables $\alpha_{ed_m} = \alpha_{ed_f} = \alpha_{sib} = 0$.

Level 2 - Allowing for Indirect Effects via Years of Schooling Years of schooling can be affected by skills acquired by age 16, investments and family circumstances. To estimate how these mediators affect the IGE both directly and indirectly via schooling, we first estimate the parameters of an equation which relates schooling to the income of the parents and each of the mediators:

$$S_i \quad = \quad \beta_C C_i + \boldsymbol{\beta}_I \boldsymbol{I}_i + \boldsymbol{\beta}_F \boldsymbol{F}_i + \beta_{Y_P} \ln Y_{Parent,i} + u_i^S \tag{6}$$

Now allowing S_i to depend on the other covariates, total differentiation of equation (3) with respect to $\ln Y_{Parent,i}$ yields:

$$\begin{aligned} \frac{d \ln Y_i}{d \ln Y_{Parent,i}} &= \alpha_S \left(\frac{\partial S_i}{\partial C_i} \frac{dC_i}{d \ln Y_{Parent,i}} + \frac{\partial S_i}{\partial \boldsymbol{I}_i} \frac{d\boldsymbol{I}_i}{d \ln Y_{Parent,i}} + \frac{\partial S_i}{\partial \boldsymbol{F}_i} \frac{d\boldsymbol{F}_i}{d \ln Y_{Parent,i}} + \beta_{Y_P} \right) \\ &\quad + \alpha_C \frac{dC_i}{d \ln Y_{Parent,i}} + \boldsymbol{\alpha}_I \frac{d\boldsymbol{I}_i}{d \ln Y_{Parent,i}} + \boldsymbol{\alpha}_F \frac{d\boldsymbol{F}_i}{d \ln Y_{Parent,i}} + \alpha_{Y_P} \end{aligned} \tag{7}$$

Noting from equation (6) that, for example, the effect of cognition on schooling is $\frac{\partial S_i}{\partial C_i} = \beta_C$, we can now compute the fraction of the IGE explained by cognition to be:

$$(\underbrace{\alpha_C \cdot \kappa_C}_{\text{Direct Effect}} + \underbrace{\alpha_S \cdot \beta_C \cdot \kappa_C}_{\text{Indirect Effect via schooling}})/\rho. \tag{8}$$

The "direct effect" of age 16 cognition on earnings is the same effect shown in level 1: $\alpha_C \cdot \kappa_C$. The new, indirect effect comes from parental income affecting schooling, through cognition: $\frac{\partial S_i}{\partial C_i} \cdot \frac{dC_i}{d \ln Y_{Parent,i}} = \beta_C \cdot \kappa_C$. As before, we know that $\frac{\partial \ln Y_i}{\partial S_i} = \alpha_S$. Together these imply that the share of the IGE explained by cognition, both directly as well as indirectly through schooling, is the expression in (8). Similar calculations are done to find the fractions explained by investments and family background.

We can now estimate whether there is any direct effect of parental income on schooling (which affects the IGE), or whether the effect is entirely mediated through the other covariates. For this, we estimate how much the coefficient on schooling differs between a regression of schooling on parental income with no other covariates (κ_S from an equation analogous to (5) but with schooling in place of cognition), versus one that includes all other covariates (β_{Y_P} from equation (6)). The fraction of the effect of parental income on schooling which is not mediated by the other covariates then becomes $\frac{\beta_{Y_P}}{\kappa_S}$. The fraction of the IGE that is explained by schooling, $\frac{\alpha_S \kappa_S}{\rho}$, can therefore be split into a part that is mediated by other covariates, $\frac{\kappa_S - \beta_{Y_P}}{\kappa_S} \cdot \frac{\alpha_S \kappa_S}{\rho}$, and the remaining part $\frac{\beta_{Y_P}}{\kappa_S} \cdot \frac{\alpha_S \kappa_S}{\rho}$.

We then test whether groups of covariates jointly have a significant effect on years of schooling. That is, in equation (6) we test for the joint significance of time investments, $\beta_{inv_7} = \beta_{inv_{11}} = \beta_{inv_{16}} = 0$, school quality, $\beta_{SQ_7} = \beta_{SQ_{11}} = \beta_{SQ_{16}} = 0$, and all family background variables $\beta_{ed_m} = \beta_{ed_f} = \beta_{sib} = 0$.

Level 3 - Allowing for Indirect Effects via Years of Schooling and Cognition In level 3 we allow investments and family background to affect cognition (C_i). We estimate the following equation:

$$C_i = \boldsymbol{\gamma}_I \boldsymbol{I}_i + \boldsymbol{\gamma}_F \boldsymbol{F}_i + \gamma_{Y_P} \ln Y_{Parent,i} + u_i^C \tag{9}$$

We model cognition as a function of all investments made up to age 16. This approach can be thought of as the reduced form to a set of structural cognition production functions of the sort estimated in Agostinelli and Wiswall (2016), Cunha and Heckman (2008) and Bolt et al. (2020), if the cognition functions at different ages are linear in their arguments.[10]

Equation (9) parsimoniously captures how all of the different channels potentially impact cognition. Allowing for the effect of parental income on cognition to be mediated by parental investments and family background we can see that we can insert $\frac{dC_i}{d\ln Y_{Parent,i}} = \frac{\partial C_i}{\partial \boldsymbol{I}_i}\frac{d\boldsymbol{I}_i}{d\ln Y_{Parent,i}} + \frac{\partial C_i}{\partial \boldsymbol{F}_i}\frac{d\boldsymbol{F}_i}{d\ln Y_{Parent,i}}$ into equation (7) to infer the total impact of investments and family background and can use this to compute the fractions of the IGE explained when allowing covariates to affect cognition and schooling. For example, the fraction of the IGE explained by time investments at age 16 is:

$$[\underbrace{\alpha_{inv_{16}}}_{\substack{\text{Direct Effect}\\ \text{of } inv_{16} \text{ on Earnings}}} + \underbrace{\beta_{inv_{16}}\alpha_S}_{\substack{\text{Indirect Effect}\\ \text{of } inv_{16} \text{ via schooling}}} + \underbrace{(\underbrace{\alpha_C}_{\substack{\text{Direct Effect}\\ \text{of cognition on Earnings}}} + \underbrace{\beta_C\alpha_S}_{\substack{\text{Indirect Effect}\\ \text{of cognition via schooling}}})\gamma_{inv_{16}}}_{\text{Indirect Effect of } inv_{16} \text{ via cognition}}] \cdot \kappa_{inv_{16}}/\rho$$

where $\kappa_{inv_{16}}$ is a regression coefficient of inv_{16} on parental income and captures the fact that richer parents invest more in their kids. We can also estimate how much of the effect of parental income on cognition can be explained by investments and family background: $\frac{\kappa_C - \gamma_{Y_P}}{\kappa_C}$.

Analogously to the previous levels, we now test whether groups of covariates jointly have a significant effect on cognition. That is, in equation (9) we test for the joint significance of time investments at all

[10] In particular, note that if the cognition production function is:

$$C_{i,t} = \varrho_t C_{i,t-1} + \Gamma_{inv,t} inv_{it} + \Gamma_{sq,t} sq_{it} + \Gamma_{ed_m,t} ed_{m,i} + \Gamma_{ed_f,t} ed_{f,i} + \Gamma_{siblings,t} siblings_i + \Gamma_{parinc,t} \ln Y_{Parent,i} + \Upsilon_{i,t}^C$$

where $t-1$ is the previous period (so if $t = 16$, then $t-1$=11), and

$$C_{i,0} = \Gamma_{ed_m,0} ed_{m,i} + \Gamma_{ed_f,0} ed_{f,i} + \Gamma_{siblings,0} siblings_i + \Gamma_{parinc,0} \ln Y_{Parent,i} + \Upsilon_{i,0}^C$$

then recursive substitution on the above equations yields:

$$\begin{aligned} C_{i,16} &= \Gamma_{inv,16} inv_{i16} + \Gamma_{sq,16} sq_{i16} + \varrho_{16}\big(\Gamma_{inv,11} inv_{i11} + \Gamma_{sq,11} sq_{i11}\big) + \varrho_{16}\varrho_{11}\big(\Gamma_{inv,7} inv_{i7} + \Gamma_{sq,7} sq_{i7}\big) \\ &+ \big(\Gamma_{ed_m,16} + \varrho_{16}\Gamma_{ed_m,11} + \varrho_{16}\varrho_{11}\Gamma_{ed_m,7} + \varrho_{16}\varrho_{11}\varrho_7\Gamma_{ed_m,0}\big) ed_{m,i} \\ &+ \big(\Gamma_{ed_f,16} + \varrho_{16}\Gamma_{ed_f,11} + \varrho_{16}\varrho_{11}\Gamma_{ed_f,7} + \varrho_{16}\varrho_{11}\varrho_7\Gamma_{ed_f,0}\big) ed_{f,i} \\ &+ \big(\Gamma_{siblings,16} + \varrho_{16}\Gamma_{siblings,11} + \varrho_{16}\varrho_{11}\Gamma_{siblings,7} + \varrho_{16}\varrho_{11}\varrho_7\Gamma_{siblings,0}\big) siblings_i \\ &+ \big(\Gamma_{parinc,16} + \varrho_{16}\Gamma_{parinc,11} + \varrho_{16}\varrho_{11}\Gamma_{parinc,7} + \varrho_{16}\varrho_{11}\varrho_7\Gamma_{parinc,0}\big) \ln Y_{Parent,i} + u_i^C \end{aligned} \tag{10}$$

where $C_i = C_{i,16}, u_i^C = \Upsilon_{i,16}^C + \varrho_{16}\Upsilon_{i,11}^C + \varrho_{16}\varrho_{11}\Upsilon_{i,7}^C + \varrho_{16}\varrho_{11}\varrho_7\Upsilon_{i,0}^C$. Note that equation (10) implies that reduced form equation (9).

ages, $\gamma_{inv_7} = \gamma_{inv_{11}} = \gamma_{inv_{16}} = 0$, school quality at all ages, $\gamma_{SQ_7} = \gamma_{SQ_{11}} = \gamma_{SQ_{16}} = 0$, and all family background variables $\gamma_{ed_m} = \gamma_{ed_f} = \gamma_{sib} = 0$.

Level 4 - Allowing for Indirect Effects via Years of Schooling, Cognition, and Investments Finally, we also allow for investments to be affected by family background. We estimate the parameters of equations that relate each element of the vector $\boldsymbol{I}$ (containing parental time and school quality investments at each age) to family background characteristics and parental income. For example, for age 16 time investments we estimate:

$$inv_{16,i} \quad = \quad \boldsymbol{\delta}_F \boldsymbol{F}_i + \delta_{Y_P} \ln Y_{Parent,i} + u_i^{inv_{16}} \tag{11}$$

The share of the IGE explained by an element of $\boldsymbol{F}$, for example, maternal education (ed_m), then becomes:

$$\Big\{ \underbrace{\alpha_{ed_m}}_{\substack{\text{Direct Effect}\\ \text{of mum ed on Earnings}}} + \underbrace{\alpha_S \beta_{ed_m}}_{\substack{\text{Indirect Effect}\\ \text{of mum ed via Schooling}}} + \underbrace{(\alpha_C + \beta_C \alpha_S)\gamma_{ed_m}}_{\substack{\text{Indirect Effect}\\ \text{of mum ed via Cognition}}} +$$
$$\underbrace{\Big[\underbrace{\alpha_{inv_{16}}}_{\substack{\text{Direct Effect}\\ \text{of } inv_{16} \text{ on Earnings}}} + \underbrace{\beta_{inv_{16}} \alpha_S}_{\substack{\text{Indirect Effect}\\ \text{of } inv_{16} \text{ via schooling}}} + \underbrace{\Big(\underbrace{\alpha_C}_{\substack{\text{Direct Effect}\\ \text{of cognition on Earnings}}} + \underbrace{\beta_C \alpha_S}_{\substack{\text{Indirect Effect}\\ \text{of cognition via schooling}}} \Big)}_{\text{Indirect Effect of } inv_{16} \text{ via cognition}} \gamma_{inv_{16}} \Big] \delta_{ed_m, inv_{16}}}_{\substack{\text{Indirect effect}\\ \text{via } inv_{16}}} \Big\} \cdot \kappa_{ed_m} / \rho$$

where κ_{ed_m} is the parental income gradient for maternal education. This captures how maternal education directly affects lifetime earnings as well as all indirect pathways running through each of years of schooling, cognition, and investments. Again, we can estimate the extent to which the relationship between investments at age 16 and parental income is mediated by family background: $\frac{\kappa_{inv_{16}} - \delta_{Y_P}}{\kappa_{inv_{16}}}$.

4.3 Latent Factors and Measurement Error Correction

Following recent developments in the literature on human capital development, we interpret the skill and investment measures in our dataset as noisy measures of unobserved, underlying factors. We allow for investments to be potentially multi-dimensional.[11] In this section we first address the question of how many underlying dimensions of investments to model. We then explain how we deal with the fact that we observe multiple, likely noisy, measures of skills and investment.

[11]In the robustness check in Section 5.5, we also allow for skills to be multidimensional.

4.3.1 Choosing the Number of Latent Factors

Given the large number of measures of investments available, it is difficult to determine a priori how many underlying latent factors there are. For example, the number of outings with parents, parental interest in the child's education, and the student-teacher ratio at the child's school may all represent measures of investment in the child's human capital. However, they may represent three distinct types of investments (e.g. leisure time with child, academic interest of parents, school environment). Our approach is to carefully select the number of underlying factors via exploratory factor analysis (Gorsuch et al. (2003), Thompson (2004), Heckman et al. (2013)), which leads us to retain only two factors which we label time investments and school quality. Further description of these procedures, tables including the eigenvalues and loadings, as well as further information on our measures can be found in Appendix D.1.

4.3.2 Using Latent Factors in Our Analysis

We now describe how our measures of investments and cognition relate to these underlying latent factors. Following previous literature (Cunha and Heckman (2008), Agostinelli and Wiswall (2016)), we assume a linear relationship between measures (Z) and underlying latent factors $\omega \in \{C, inv, sq\}$:

$$Z_{\omega,i,t,j} = \mu_{\omega,t,j} + \lambda_{\omega,t,j}\omega_{i,t} + \epsilon_{\omega,i,t,j}$$

Here, $Z_{\omega,i,t,j}$ denotes measure j of latent factor ω (e.g. a math score as a measure of latent cognitive skills) for individual i at time t. $\mu_{\omega,t,j}$ and $\lambda_{\omega,t,j}$, respectively, are the location and scale of this measure and are constant across individuals. $\epsilon_{\omega,i,t,j}$ denotes an idiosyncratic measurement error, assumed to be independent across individuals, measures, and time. The measurement errors are also assumed to be independent of the latent variables and all other controls and shocks. As the latent factors do not have a natural scale or location, we normalize their means to be zero in every period, and their variances to be one. This allows us to estimate the location and scale parameter for each measure (see Appendix D.2 for details). We then predict the latent factors for each individual, using the Bartlett score method (Heckman et al. (2013)). Bartlett scores are a linear combination of all retained measures, inversely weighted by their noise. This means that measures with little measurement error get more weight than those with a lot of measurement error. The measurement error in each cognitive skill and investment measure is minimized, but not eliminated, by combining them via the Barlett scoring method. As a result, we cannot simply estimate the above equations via OLS, but have to correct for the remaining measurement error. We do this by using the errors-in-variables correction as suggested by Heckman et al. (2013). Appendix D.2.4 gives more details.

5 Results

5.1 Estimates of the IGE

Applying the approach described in Section 4.1, we first present our estimates of the the IGE. Table 3 shows estimates both with and without corrections for measurement error. Without correcting for measurement error, the estimated IGE is 0.16 and 0.12 for males and females respectively. These estimates are similar to estimates in other studies using the same data (Belfield et al. (2017), Gregg et al. (2016)), and are robust to alternative measures of earnings (see Appendix C). This increases to 0.32 and 0.24 once we account for measurement error that comes from using parental income at age 16 as a proxy for parental lifetime income.

Table 3: IGE estimates

	Male	Female	Male uncorrected	Female uncorrected
IGE	0.317	0.236	0.155	0.115
	(0.097)	(0.105)	(0.045)	(0.050)
N	1350	1347	1350	1347

Note: The first two columns report the IGE estimates for male and female children, after parental income has been corrected for measurement error. The third and fourth column report estimates without this correction.

5.2 Mediation Analysis

Table 4 gives the full results of our mediation analysis. Overall, variation in the factors we consider – schooling, cognition, investments and family background – can explain 54% (62%) of the intergenerational elasticity in earnings of males (females). The next four subsections discuss, in turn, the results in each of the four columns of Table 4, which represent the fractions of the IGE that can explained by each covariate at each of our four levels of mediation.

Baseline - Direct effects on lifetime earnings The baseline column in Table 4 decomposes the IGE by only allowing for direct effects of each factor on the IGE. We find that differences in cognition and years of schooling explain the largest fraction of the IGE. The fraction explained by these two factors is high for both males and females, but their relative importance differs across genders. For males, 33% of the IGE is explained by cognition and 10% by years of schooling. For females, 13% are explained by cognition and 43% by years of schooling.

We find that parental investments have only a small impact, and family background has virtually no impact on lifetime earnings once we control for cognition, school attainment, and parental income, as we

Table 4: Multilevel Mediation Analysis: Share of IGE Explained

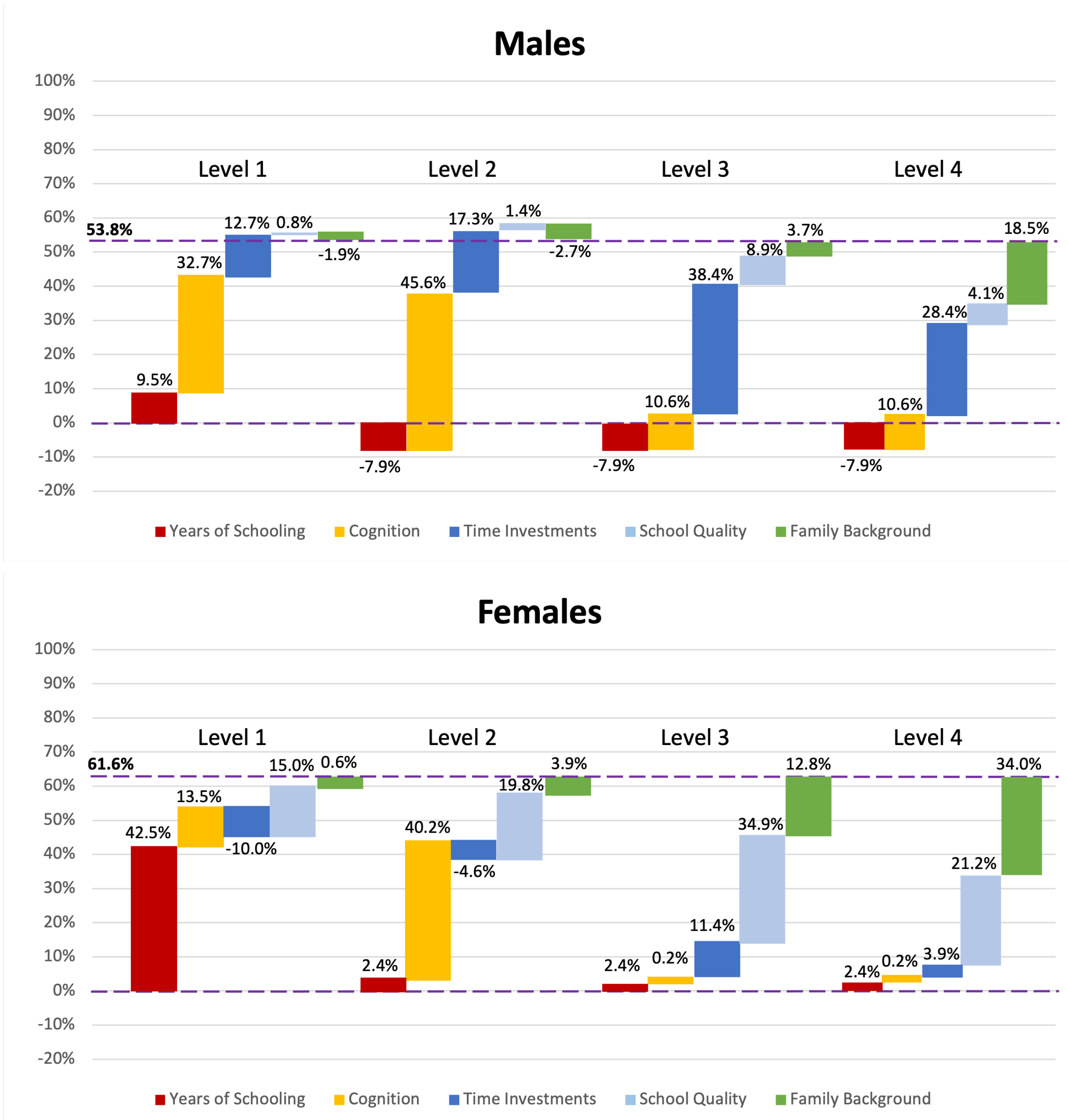
	Males				**Females**			
	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4
Years of Schooling	**0.095**	**-0.079**	**-0.079**	**-0.079**	**0.425**	0.024	0.024	0.024
	[0.020, 0.191]	[-0.189, -0.005]	[-0.189, -0.005]	[-0.189, -0.005]	[0.165, 1.361]	[-0.249, 0.340]	[-0.249, 0.340]	[-0.249, 0.340]
Cognition	**0.327**	**0.456**	0.106	0.106	**0.135**	**0.402**	0.002	0.002
	[0.164, 0.652]	[0.258, 0.931]	[-0.069, 0.329]	[-0.069, 0.329]	[0.007, 0.508]	[0.185, 1.237]	[-0.286, 0.204]	[-0.286, 0.204]
Investments	0.135	0.187	**0.473**	**0.325**	0.050	0.151	**0.463**	0.251
	[-0.138, 0.393]	[-0.087, 0.480]	[0.175, 0.930]	[0.084, 0.643]	[-0.271, 0.302]	[-0.142, 0.560]	[0.126, 1.527]	[-0.049, 0.947]
Time Investments	0.127	0.173	**0.384**	**0.284**	-0.100	-0.046	0.114	0.039
	[-0.090, 0.331]	[-0.029, 0.431]	[0.149, 0.791]	[0.085, 0.589]	[-0.428, 0.083]	[-0.372, 0.157]	[-0.106, 0.655]	[-0.223, 0.325]
... Age 7	0.126	0.147	0.143	0.111	0.143	0.157	**0.176**	0.105
	[-0.080, 0.441]	[-0.059, 0.450]	[-0.075, 0.441]	[-0.067, 0.325]	[-0.010, 0.690]	[-0.005, 0.669]	[0.002, 0.824]	[-0.006, 0.458]
... Age 11	-0.054	-0.027	0.076	0.066	**-0.180**	**-0.175**	-0.133	-0.083
	[-0.392, 0.183]	[-0.341, 0.223]	[-0.162, 0.386]	[-0.147, 0.326]	[-0.705, -0.017]	[-0.702, -0.019]	[-0.544, 0.037]	[-0.385, 0.034]
... Age 16	0.056	0.053	**0.166**	**0.108**	-0.062	-0.029	**0.070**	0.016
	[-0.008, 0.146]	[-0.010, 0.147]	[0.080, 0.304]	[0.030, 0.205]	[-0.285, 0.012]	[-0.156, 0.054]	[0.008, 0.289]	[-0.055, 0.137]
School Quality	0.008	0.014	0.089	0.041	**0.150**	**0.198**	**0.349**	**0.212**
	[-0.124, 0.135]	[-0.121, 0.142]	[-0.076, 0.234]	[-0.106, 0.126]	[0.008, 0.483]	[0.038, 0.698]	[0.139, 1.187]	[0.014, 0.619]
... Age 7	-0.001	-0.001	0.000	-0.000	0.047	0.044	0.046	0.032
	[-0.028, 0.021]	[-0.029, 0.019]	[-0.020, 0.020]	[-0.029, 0.022]	[-0.016, 0.219]	[-0.021, 0.213]	[-0.020, 0.228]	[-0.044, 0.187]
... Age 11	-0.024	-0.023	-0.044	-0.030	0.019	0.022	0.016	-0.010
	[-0.137, 0.075]	[-0.139, 0.080]	[-0.179, 0.051]	[-0.135, 0.026]	[-0.061, 0.109]	[-0.064, 0.098]	[-0.074, 0.080]	[-0.129, 0.030]
... Age 16	0.033	0.038	**0.133**	0.072	0.084	**0.132**	**0.287**	**0.191**
	[-0.017, 0.085]	[-0.016, 0.094]	[0.048, 0.231]	[-0.006, 0.140]	[-0.020, 0.300]	[0.021, 0.408]	[0.102, 0.851]	[0.043, 0.591]
Family Background	-0.019	-0.027	0.037	**0.185**	0.006	0.039	0.128	**0.340**
	[-0.131, 0.089]	[-0.146, 0.078]	[-0.089, 0.166]	[0.063, 0.453]	[-0.288, 0.221]	[-0.220, 0.323]	[-0.089, 0.554]	[0.136, 1.174]
Mother's education	-0.051	-0.049	-0.029	0.020	-0.043	-0.024	0.010	0.104
	[-0.137, 0.012]	[-0.139, 0.015]	[-0.103, 0.038]	[-0.039, 0.095]	[-0.304, 0.152]	[-0.300, 0.173]	[-0.230, 0.262]	[-0.047, 0.567]
Father's education	0.016	0.008	0.035	0.084	0.068	0.081	0.126	**0.227**
	[-0.077, 0.096]	[-0.084, 0.088]	[-0.063, 0.114]	[-0.008, 0.184]	[-0.088, 0.277]	[-0.079, 0.348]	[-0.018, 0.418]	[0.076, 0.679]
Number of Siblings	0.016	0.014	0.031	**0.081**	-0.019	-0.019	-0.008	0.009
	[-0.012, 0.053]	[-0.015, 0.047]	[-0.001, 0.085]	[0.008, 0.191]	[-0.140, 0.027]	[-0.146, 0.028]	[-0.099, 0.023]	[-0.033, 0.063]
Total	0.538	0.538	0.538	0.538	0.616	0.616	0.616	0.616
N	1350	1350	1350	1350	1347	1347	1347	1347

Notes: 95% Confidence intervals constructed using 250 clustered bootstrap replications. Coefficients that are significant at the 5% level are **bold**. The row labelled 'Investments' shows the fraction explained by the sum of all investments, i.e. time investments and school quality. The row labelled 'Time Investments' shows the fraction explained by the sum of all time investments, i.e. for ages 7, 11, 16. Similarly for 'School Quality' and 'Family Background'.

do at this stage of mediation. This is an attractive aspect of both our data and our framework. Whereas most of the literature *assumes* that time and school quality investments impact lifetime earnings only through their role in promoting higher cognition and school attainment, we formally test to see if that is the case. We find that neither parental education nor number of siblings play a role, conditional on cognition and years of schooling. Hence, conditional on parents' income, we find no evidence that parents transmit an advantage to their children's earnings beyond that which comes from improved cognition and schooling outcomes.

These estimated shares are obtained using estimates of the relationship between lifetime earnings and our covariates in equation (3) and the relationship between parental income and those covariates

Figure 2: Graphical Summary of Mediation Analysis: Share of IGE Explained



(equation (5)). Table 5 shows determinants of lifetime earnings, schooling and cognition for both males and females. Columns 1 and 2 – the relevant ones for understanding the baseline mediation results – show that cognition and years of schooling are important determinants of lifetime earnings. For men, one more year of schooling increases lifetime earnings by 5.7%, whereas a one standard deviation increase in cognition increases lifetime earnings by 14%. For women, we find a larger return to education and a smaller return to cognition compared to males. The bottom of Table 5 reports p-values for the tests of

Table 5: Determinants of Lifetime Earnings, Schooling, Cognition

	Lifetime Earnings		**Years of Schooling**		**Cognition**	
	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)
Years of Schooling	**0.057**	**0.134**				
	(0.013)	(0.014)				
Cognition	**0.139**	0.064	**0.973**	**0.944**		
	(0.030)	(0.036)	(0.054)	(0.062)		
Time Investments						
Age 7	0.041	0.062	0.124	0.047	-0.006	0.043
	(0.040)	(0.043)	(0.090)	(0.087)	(0.066)	(0.069)
Age 11	-0.015	-0.076	0.128	0.016	0.142	0.094
	(0.045)	(0.041)	(0.101)	(0.078)	(0.091)	(0.088)
Age 16	0.032	-0.042	-0.024	**0.169**	**0.328**	**0.349**
	(0.023)	(0.027)	(0.053)	(0.052)	(0.036)	(0.032)
School Quality Investments						
Age 7	-0.012	0.047	0.006	-0.019	0.078	0.008
	(0.060)	(0.025)	(0.169)	(0.047)	(0.056)	(0.029)
Age 11	-0.013	0.025	0.011	0.024	-0.057	-0.037
	(0.034)	(0.030)	(0.053)	(0.136)	(0.038)	(0.068)
Age 16	0.024	0.035	0.071	**0.151**	**0.362**	**0.342**
	(0.022)	(0.023)	(0.050)	(0.049)	(0.029)	(0.032)
Family Background						
Mother's education	-0.025	-0.010	0.020	0.031	**0.051**	0.041
	(0.020)	(0.023)	(0.041)	(0.055)	(0.024)	(0.025)
Father's education	0.007	0.015	-0.062	0.022	**0.057**	**0.050**
	(0.020)	(0.017)	(0.038)	(0.036)	(0.025)	(0.022)
Number of siblings	-0.013	0.025	0.030	-0.001	**-0.070**	**-0.072**
	(0.013)	(0.016)	(0.029)	(0.032)	(0.021)	(0.020)
Log parental income	0.146	0.091	-0.440	0.042	0.173	0.002
	(0.110)	(0.111)	(0.252)	(0.236)	(0.165)	(0.124)
P-values for joint significance:						
Time Investments	0.708	0.842	0.490	0.315	0.096	0.031
School Quality	0.501	0.285	0.424	0.183	0.017	0.009
Family Background	0.291	0.276	0.218	0.408	0.012	0.020
N	1350	1347	1350	1347	1350	1347

Note: Lower panel shows p-values of F-tests for joint significance of time investments at all ages, school quality at all ages, and all family background variables. The null hypothesis is that the relevant group of variables jointly have zero coefficient in the corresponding regression. Importantly, the F-tests test restrictions while accounting for measurement error in the data. See Appendix (E) for more details.

the joint significance of time investments at different ages, school quality at different ages, and family background. These are modified F-tests that account for measurement error in all the variables we use – see Appendix E for details. We find that the null hypothesis of no effect of these categories on lifetime earnings cannot be rejected. This confirms the usual view that parental time investments and family background affect lifetime earnings primarily through the channel of cognition and schooling. Estimates for the relationship between mediators and parental income (equation (5)) can be found in Appendix F Table 13.

Level 2 - Allowing for Indirect Effects via Years of Schooling The second column in each of Table 4's two panels shows results which allow the fraction of the IGE explained by years of schooling to be mediated by the other covariates – cognition, investments and family background. Once we allow for this mediation, the fraction of the IGE explained by years of schooling is no longer statistically different from zero for women and becomes negative for men. For both males and females, the higher educational attainment of individuals from high income households is completely mediated by these three factors, as opposed to being a direct effect of parental income on years of schooling. The most important mediator is cognition, which is consistent with the literature. Multiple previous studies have shown that those with greater cognitive skills are more likely to complete more years of education, in part due to the greater success they appear to have in higher education (Keane and Wolpin (2001), Carneiro and Heckman (2002), Arcidiacono (2005)). For males, we find that cognition, in addition to explaining 33% of the IGE through a direct effect, also accounts for 13% of the IGE through an indirect effect on schooling, thus explaining a total share of 46% of the IGE. The total effect for women is similar - at 40%. Neither the education of parents nor the number of siblings has a significant impact on years of school attainment, once cognition and investments are controlled for.

The estimates in Columns 3 and 4 of Table 5, which show the determinants of schooling, help to understand the drivers of this result. Cognition is the major determinant of educational attainment - a one standard deviation increase in cognition at age 16 leads to almost a whole additional year of schooling. This estimate is remarkably similar for both males and females.

The F-tests reported at the bottom of the table show that we cannot reject the null hypothesis that time investments, school quality, and family background have no effect on years of schooling, once cognition is controlled for. This highlights the importance of age 16 cognition as a key explanatory variable for lifetime earnings, as it has both an important direct effect on earnings, and also a significant indirect effect through schooling.

Level 3 - Allowing for Indirect Effects via Years of Schooling and Cognition In Level 3, the results of which are reported in the third column of each of Table 4's two panels, we allow the effect of parental income on age 16 cognition to be mediated by parental investments and family background. We find that once we consider indirect effects of investments on cognition and schooling, investments become the key driver of the IGE. Over 40% of the IGE can be explained by the increased time and school quality investments of high income parents, mostly through their impact on cognition. The relative importance of different types of investment varies by gender – time investments are quantitatively more important than school quality for men, whereas for women this is reversed. Both types of investment are important, with time investments more important than school quality early in childhood, and school quality investments becoming relatively more important later in childhood. For example, for men, parental time investments at age 7 contribute much more than school quality investments at age 7 to the IGE (14% versus 0%). However, parental time and school quality investments at age 16 contribute similar amounts to the IGE (17% and 13%, respectively). All investments appear important at age 16, which is of particular interest given the recent interest in early life investments. Interestingly, parental education and number of siblings have only a modest impact on cognition once we control for investments, and thus the shares of the IGE explained by these factors only change slightly between level 2 and level 3.

Table 5 presents the drivers of these results. The third panel, which shows the relationship between cognition and each of investments and background, indicates that time investments and school quality have strong effects on cognition at age 16. This result is in keeping with the results from the large recent literature on the estimation of production functions for child ability (Cunha and Heckman (2008), Cunha et al. (2010), Attanasio et al. (2020)).

The education of the mother and the father have a small positive effect on cognition even after controlling for investments. This could be due to direct ability transmission which we cannot capture, or due to differential productivity of higher education parents in producing child ability. We also find a negative effect of the presence of siblings.

Level 4 - Allowing for Indirect Effects via Years of Schooling, Cognition, and Investments In our final mediation stage we additionally allow the effect of parental income on investments to be mediated by measures of family background – parental education and number of siblings. Of these, parental education plays an important role, especially for females, in mediating the effect of investments. Although in our level 3 analysis we find that parental education has little effect once we control for investments, in level 4 we find that parental education is important because it impacts investments. For females, mother's education explains 1% of the IGE when we control for investments, but explains 10% of the IGE in level 4 when we allow investments to be a function of mother's education.

However, the investments of high income parents cannot be fully explained by parental education or by the number of siblings. In Table 4, the fraction of the IGE explained by investments at level 3 is 0.473 (0.463), yet it remains at 0.325 (0.251) for men (women) once we allow for family background as a mediator in level 4. This suggests a potential direct role for parents' income in determining investments. Turning to Table 6, which provides estimates of our investment equation (11), we can see strong parental income gradients for time investments at younger ages and for school quality at older ages after controlling for family background. Together with the large impacts of investments on cognition and years of schooling (Table 5), we conclude that parental income gradients in investments are a powerful source for perpetuating inequality.

Table 6: Determinants of Parental Investments

Time Investments

	Males			Females		
	Age 7	Age 11	Age 16	Age 7	Age 11	Age 16
Mother's Education	**0.080**	0.064	**0.112**	**0.095**	**0.110**	**0.182**
	(0.037)	(0.039)	(0.026)	(0.030)	(0.039)	(0.028)
Father's Education	0.058	0.024	**0.090**	**0.068**	0.055	0.045
	(0.031)	(0.038)	(0.025)	(0.027)	(0.035)	(0.025)
Number of Siblings	**-0.310**	**-0.237**	**-0.136**	**-0.261**	**-0.205**	**-0.175**
	(0.022)	(0.024)	(0.018)	(0.025)	(0.028)	(0.020)
Log parental income	**0.746**	**1.021**	**0.362**	0.323	0.350	0.081
	(0.210)	(0.205)	(0.148)	(0.178)	(0.188)	(0.161)
N	1350	1350	1350	1347	1347	1347

School Quality Investments

	Males			Females		
	Age 7	Age 11	Age 16	Age 7	Age 11	Age 16
Mother's Education	0.052	0.058	**0.104**	0.023	**0.169**	0.056
	(0.039)	(0.049)	(0.033)	(0.086)	(0.067)	(0.034)
Father's Education	0.004	**0.171**	**0.124**	0.035	**0.103**	**0.116**
	(0.039)	(0.056)	(0.036)	(0.047)	(0.051)	(0.027)
Number of Siblings	-0.034	**-0.052**	**-0.084**	-0.056	-0.026	-0.023
	(0.026)	(0.019)	(0.017)	(0.051)	(0.016)	(0.021)
Log parental income	-0.032	0.410	0.231	0.165	-0.113	**0.374**
	(0.178)	(0.217)	(0.141)	(0.314)	(0.237)	(0.172)
N	1350	1350	1350	1347	1347	1347

Note: These tables show the regression of investments on family background variables. This provides some explanation of the difference between levels 3 and 4 of our mediation analysis. All investment measures here are corrected for measurement error.

Summary Figure 2 summarizes the main results of our mediation analysis. The most striking aspect of the graph for men is the substantial role that cognition plays in explaining the IGE in the baseline, with a comparatively modest impact of family background. This role for cognition is substantially attenuated by level 3 – that is, the seeds for the cognitive advantage enjoyed by the children of richer parents are found in the greater investments that those parents make. For women, we see that cognition and schooling are the most important in the baseline, but by level 4, when these are allowed to be mediated by investments and family background, we find that those latter factors play an important role.

In the next section we provide a summary and compare our results to the literature.

5.3 Key Findings and Comparisons to the Literature

In this section, we summarize our main findings and relate them to some important results in the literature.

1) Lifetime earnings are mainly driven by cognition and years of schooling. Parental investments and family background have no significant effect on the IGE once we control for cognition and years of schooling.

Columns 1 and 2 of Table 5 shows that cognition and schooling have sizeable and significant effects on lifetime earnings for men and women, respectively. The return to a year of schooling on lifetime earnings is 5.7% (13.4%) for men (women), which falls within the range of commonly reported returns to schooling, as summarized in Card (1999). For males, the effect of a standard deviation of cognition on lifetime earnings is 13.9%, similar to the estimates in Heckman et al. (2006). Whilst many papers *assume* that earnings is mainly determined by schooling and cognition at the beginning of adulthood (Huggett et al. (2011), Lee and Seshadri (2019), Daruich (2018)), our analysis confirms that once schooling and cognition are controlled for, other early life factors have no significant effect.

2) The role of years of schooling in explaining the relationship between parents' income and children's earnings is mediated entirely by differences in cognition. This means that direct effects of parental income on schooling are insignificant in our setting.

Our paper contributes to the important question as to whether borrowing constraints limit educational attainment of those from low income families and thus propagate inequality. Consistent with Arcidiacono (2005), Stinebrickner and Stinebrickner (2008), Carneiro and Heckman (2002) and Carneiro et al. (2011), we find a large effect of cognitive skills on schooling decisions. Our finding that parental income has no significant impact on educational attainment, conditional on cognition, is consistent with Belley and Lochner (2007), who find that conditional on ability, parental income did not matter for the schooling decisions of the NLSY79 cohort (although they find that it does for more recent cohorts). Part of the explanation for this may lie in the fact that there were no university tuition fees in the UK for our cohort.

We do not find a direct effect of parental education on years of schooling, once we control for cognition. This can be interpreted as evidence for the intergenerational persistence in education being transmitted mostly through cognition (Lee and Seshadri (2019)).

3) Parental investments in time and school quality explain most of the differences in cognition in individuals from low and high income families.

Consistent with recent literature, we find that parental time investments are crucial for child human capital development (see Heckman and Mosso (2014) for a summary). Whilst we cannot rule out potential endogeneity of our investments, several papers have found parents compensate for shocks during childhood (Attanasio et al. (2017), Attanasio et al. (2020)) – that is, parents invest more in children who have lower ability. If this dynamic is at work in our cohort, our estimates of the effect of time investments on cognition will be downward biased and our results regarding the impact of time investments on the IGE would be a lower bound.

Several papers estimate the effect of school quality on lifetime outcomes; the majority use a single measure such as classroom size (Akerhielm (1995), Altonji and Dunn (1996), Dearden et al. (2002), Chetty et al. (2011)). Our approach has the benefit of combining multiple measures of school quality and correcting for measurement error. This leads us to find larger effects of school quality on cognition than, for example, Akerhielm (1995). In Appendix H, we provide further comparisons regarding the effect of school quality on lifetime earnings.

4) We find that income-related differences in child investments (and their consequences) explain 32.5% of the IGE, whereas family background-related differences in child investments account for 18.5% of the IGE. This means that parental income directly affects investments, controlling for family background.

The importance of family circumstances during childhood on human capital investments has recently been documented in the literature. Carneiro et al. (2013) estimate the causal effect of an increase in maternal education on reading to the child at ages 7-8 to be approximately 6% of a standard deviation. This is similar in magnitude to the effect of maternal education on investments which we find at those ages (8-10%). Bhalotra and Clarke (2020) find that an exogenous increase in the number of children in the household due to twin birth decreases the amount that parents read to their children. Our results point in the same direction.

Our finding that it is not only family background, but parental income itself which matters for parental investments is consistent with the growing literature which studies the effect of financial transfers on human capital development (Duncan et al. (2011), Milligan and Stabile (2011), Dahl and Lochner (2012)), and in particular with Caucutt and Lochner (2020), who provide evidence that the mechanism behind this link is that parental income impacts investments in children.

See Appendix H for further comparisons regarding the effect of family size and parental education on lifetime earnings.

5.4 The Importance of Correcting for Measurement Error

Table 7: Main Decomposition without Measurement Error Correction

	Males				**Females**			
	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4
Years of School	**0.177**	0.007	0.007	0.007	**0.530**	0.104	0.104	0.104
	[0.059, 0.439]	[-0.149, 0.109]	[-0.149, 0.109]	[-0.149, 0.109]	[0.234, 1.980]	[-0.234, 0.584]	[-0.234, 0.584]	[-0.234, 0.584]
Cognition	**0.175**	**0.294**	0.123	0.123	0.042	**0.282**	**0.132**	**0.132**
	[0.087, 0.468]	[0.157, 0.745]	[-0.014, 0.351]	[-0.014, 0.351]	[-0.079, 0.198]	[0.110, 1.007]	[0.010, 0.616]	[0.010, 0.616]
Investments	**0.132**	**0.175**	**0.287**	**0.178**	0.044	**0.153**	**0.243**	**0.136**
	[0.004, 0.382]	[0.042, 0.483]	[0.111, 0.732]	[0.044, 0.493]	[-0.121, 0.441]	[0.010, 0.791]	[0.083, 1.094]	[0.005, 0.702]
Family Background	0.018	0.027	0.085	**0.194**	0.009	0.087	0.146	**0.254**
	[-0.124, 0.168]	[-0.115, 0.191]	[-0.049, 0.315]	[0.029, 0.563]	[-0.268, 0.299]	[-0.136, 0.515]	[-0.057, 0.686]	[0.053, 1.016]
Total	0.502	0.502	0.502	0.502	0.626	0.626	0.626	0.626
N	1092	1092	1092	1092	1127	1127	1127	1127

Note: This replicates the mediation analysis in Table 4, except it uses an individual measure rather than latent factors for each of cognition, time investments and school quality; and we do not correct for measurement error.

Throughout this analysis, we have confronted the fact that parental income, cognition, and time and school quality investments are all measured with error. In this section we discuss the importance of allowing for measurement error in these variables. Table 7 gives results of our mediation analysis when *not* accounting for measurement error. To run this analysis, instead of using the Bartlett scores for each of factors, we choose the single measure with the highest signal-to-noise ratio for cognition, school quality, and cognition. We then conduct the same mediation analysis as outlined in Section 4, but without using the errors-in-variables approach to correct for measurement error.

Although the overall share of the IGE explained is not substantially affected by controlling for measurement error, there are two key differences between the results in Table 7 and those we reported in Table 4. First, consistent with the usual attenuating role of measurement error, the estimated coefficients on the potentially error-ridden variables are smaller when not accounting for measurement error, leading to a smaller share explained by our variables of interest. Second, the objects not measured with error explain a larger share of the IGE when not accounting for measurement error. This is as expected. Since all the variables considered are correlated with one another, erroneously understating the impact of one variable will lead us to erroneously overstate the role of the other variables.

5.5 Robustness Checks

In this section, we summarize the results of the following robustness checks: including more family background variables, including non-cognitive skills, and allowing for interactions between lifetime earnings and cognition.

Additional Family Background Variables We have focused on key drivers of lifetime earnings that are highlighted in the literature. We find that our results are robust to the inclusion of other variables that are commonly thought to be important for child development, such as marital stability, and parental age. Results are presented in Table 14 in Appendix F. The coefficients on variables included in the original specification change only very modestly. Moreover, none of the additional family background variables are found to explain a significant fraction of the IGE.

Non-cognitive skills Our approach focuses on the importance of cognition and years of schooling in explaining lifetime earnings. A number of recent papers have highlighted the importance of non-cognitive skills (e.g., Heckman et al. (2006), Papageorge et al. (2019), Todd and Zhang (2020)). To assess the importance of non-cognitive skills in our data, we use the same latent variables approach to measure non-cognitive skills that we use to measure cognition. Following Heckman et al. (2013), we use exploratory factor analysis to determine the appropriate measures to use. Previous papers using the NCDS often exploit teacher-reported behaviors at age 11 as measures of non-cognitive skills (e.g., Papageorge et al. (2019), Blanden et al. (2007)). However, these are not available at age 16. In order to allow for non-cognitive skills to enter into the mediation analysis in a manner analogous to how cognition enters, we thus focus on parent-reported behaviors at age 16 as measures of non-cognitive skills. Using exploratory factor analysis, we retain the following seven out of 14 possible measures for our non-cognitive skills factor: difficulty concentrating, destructive behavior, tearfulness, restlessness, belligerence, irritability, and disobedience. Appendix D.2.4 contains a full list of measures and results of the exploratory factor analysis.

Table 15 presents the results of the mediation analysis when non-cognitive skills are included. Non-cognitive skills explain a small and insignificant fraction of the IGE. This holds at the baseline level, but also when we allow for non-cognitive skills to potentially affect schooling. The fractions explained by years of schooling and cognition only marginally decrease when non-cognitive skills are included.

The robustness of our mediation results to the inclusion of non-cognitive skills can be explained by Table 16 in Appendix G.2, which shows the determinants of lifetime earnings and years of schooling. The coefficients on non-cognitive skills are statistically insignificant, except for in the females' years of schooling regression, and everywhere are about five times smaller than the coefficients on cognition. For example, a one standard deviation increase in cognition increases men's lifetime earnings by 13.5%, while a standard deviation increase in non-cognitive skills increases their lifetime earnings by 3.8%. Further, comparing the lifetime earnings regressions in Table 16, which control for non-cognitive ability, with our main results in Table 5, we find the effect of cognition on lifetime earnings and years of schooling is essentially unchanged, for both males and females.

Finally, and crucially for this paper, we find that the correlation of parental income with non-cognitive skills is small, especially when compared to the relationship between parental income and cognitive skills. Thus, even if non-cognitive skills were important for lifetime earnings, they would not help explain the IGE in our data. For example, Table 17 in Appendix F shows that a 1% increase in parental income is associated with a 0.5% of a standard deviation increase in cognition for females at age 16, while it is associated with only a 0.001% of a standard deviation increase in non-cognitive skills at the same age.

In short, our approach is intentionally parsimonious. We focus on what we have found to be the two most robust measures for predicting lifetime earnings (cognition and years of schooling), and we use measures of investment and family background that are important for determining these.

Complementarity between cognition and years of schooling Throughout we have used equations that are linear in factors. The linear specification facilitates easily interpretable decompositions, especially when using multiple levels. However, Bolt et al. (2020) find evidence of complementarity between schooling and cognition in the wage equation using NCDS data. Therefore we conduct a robustness check in which we allow for an interaction between cognition and schooling in the lifetime earnings equation. We find no major change in the decomposition when allowing for this complementarity. Moreover, the interaction term explains an insignificant fraction of the IGE. See Appendix G.3 for further details.

6 Conclusion

In this paper, we quantify the direct and indirect effects of a large number of potential drivers of the IGE by performing a systematic multilevel mediation analysis. Overall, we are able to explain 54% of the IGE for males and 62% for females.

The four mechanisms we model as generating a link between income of parents and children are: years of schooling, cognition of the children, investments the parents make in the children, and family background. Of these, cognition and years of schooling play the most important role when all mechanisms are considered to have only a direct effect on transmission of income across generations. Once we allow for investments at earlier ages to affect cognition and years of schooling at later ages, the perspective changes. The role of years of schooling in explaining the relationship between parents' income and children's earnings is completely mediated by differences in cognition, and the role of cognition is substantially mediated by differences in time investments. Finally, around one third of the differences in investments by rich and poor parents can be attributed to differences in family background. Once we account for the effect of family background and early investments on cognition and years of schooling, we find that family background and investments are by far the most important mechanisms explaining the IGE. Failure to account for measurement error leads to an underestimated role of investments.

Apart from these results being interesting in their own right, we believe that the methodology used in this paper provides a systematic way in which researchers can explore how potentially relevant channels operate, a step which can complement the explicit modeling of decision-making.

Our paper focuses on the link between early life circumstances and lifetime earnings through their impact on human capital. However, we do not address the channels by which human capital impacts earnings in adulthood, whether it be through higher wages or higher employment rates. We also do not consider other important adult outcomes, such as marital choices. Building on the current mediation framework for further investigations of these issues would be a fruitful direction for future research.

References

Agostinelli, F. and G. Sorrenti (2018). Money vs. Time: Family income, maternal labor supply, and child development. *University of Zurich, Department of Economics, Working Paper* (273).

Agostinelli, F. and M. Wiswall (2016, July). Identification of dynamic latent factor models: The implications of re-normalization in a model of child development. Working Paper 22441, National Bureau of Economic Research.

Aizer, A. and F. Cunha (2012, September). The production of human capital: Endowments, investments and fertility. Working Paper 18429, National Bureau of Economic Research.

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review 14*(3), 229–241.

Altonji, J. G. and T. A. Dunn (1996). Using siblings to estimate the effect of school quality on wages. *The Review of Economics and Statistics 78*(4), 365–671.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), 773–824.

Arcidiacono, P. (2005). Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica 73*(5), 1477–1524.

Attanasio, O., C. Meghir, and E. Nix (2020). Human capital development and parental investment in India. *The Review of Economic Studies 87*(6), 2511–2541.

Attanasio, O., C. Meghir, E. Nix, and F. Salvati (2017). Human capital growth and poverty: Evidence from Ethiopia and Peru. *Review of economic dynamics 25*, 234–259.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2020, January). Women's empowerment in action: Evidence from a randomized control trial in Africa. *American Economic Journal: Applied Economics 12*(1), 210–59.

Belfield, C., C. Crawford, E. Greaves, P. Gregg, and L. Macmillan (2017). Intergenerational income persistence within families. Technical report, IFS Working Papers.

Belley, P. and L. Lochner (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human capital 1*(1), 37–89.

Bhalotra, S. and D. Clarke (2020). The twin instrument: Fertility and human capital investment. *Journal of the European Economic Association 18*(6), 3090–3139.

Black, S. E., P. J. Devereux, et al. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics 4*, 1487–1541.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Blanden, J., P. Gregg, and L. Macmillan (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal 117*(519), C43–C60.

Blanden, J., P. Gregg, and L. Macmillan (2013). Intergenerational persistence in income and social class: The effect of within-group inequality. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(2), 541–563.

Bolt, U., E. French, J. H. Maccuish, and C. O'Dea (2020). Intergenerational altruism and transfers of time and money: A lifecycle perspective. Technical report, Mimeo.

Card, D. (1999). Chapter 30 - The causal effect of education on earnings. *Handbook of Labor Economics 3*, 1801–1863.

Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal 112*(482), 705–734.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–81.

Carneiro, P., C. Meghir, and M. Parey (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association 11*(suppl_1), 123–160.

Caucutt, E. M. and L. Lochner (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy 128*(3), 1065–1147.

Centre for Longitudinal Studies (2017). National child development study [data collection]. `http://doi.org/10.5255/UKDA-SN-7669-1`.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly journal of economics 126*(4), 1593–1660.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity: The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics 129*(4), 1553–1623.

Conti, G., J. J. Heckman, and R. Pinto (2016). The effects of two influential early childhood interventions on health and healthy behaviour. *The Economic Journal 126*(596), F28–F65.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*(3), 79–102.

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

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

Dahl, G. B. and L. Lochner (2012). The impact of family income on child achievement: Evidence from the Earned Income Tax Credit. *American Economic Review 102*(5), 1927–56.

Daruich, D. (2018). The macroeconomic consequences of early childhood development policies. *FRB St. Louis Working Paper* (2018-29).

Daruich, D. and J. Kozlowski (2020). Explaining intergenerational mobility: The role of fertility and family transfers. *Review of Economic Dynamics 36*, 220–245.

Dearden, L., J. Ferri, and C. Meghir (2002). The effect of school quality on educational attainment and wages. *The Review of Economics and Statistics 84*(1), 1–20.

Dearden, L., S. Machin, and H. Reed (1997). Intergenerational mobility in Britain. *The Economic Journal*, 47–66.

Deckers, T., A. Falk, F. Kosse, P. Pinger, and H. Schildberg-Hörisch (2017). Socio-economic status and inequalities in children's IQ and economic preferences.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *The Review of Economic Studies 81*(1), 137.

Duncan, G. J., P. A. Morris, and C. Rodrigues (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental psychology 47*(5), 1263.

Gayle, G.-L., L. Golan, and M. Soytas (2015). What accounts for the racial gap in time allocation and intergenerational transmission of human capital?

Gayle, G.-L., L. Golan, and M. Soytas (2018). What is the source of the intergenerational correlation in earnings? Technical report, Mimeo.

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Gorsuch, R. L., J. Shinka, and W. Velicer (2003). Handbook of psychology: Vol 2. research methods in psychology. *ed. JA Schinka and WF Velicer, Chapter Factor Analysis*, 143–164.

Gregg, P., L. Macmillan, and C. Vittori (2016). Moving towards estimating sons lifetime intergenerational economic mobility. *Oxford Bulletin of Economics and Statistics, forthcoming*.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *Journal of Economic perspectives 22*(3), 23–46.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review 96*(4), 1308–1320.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

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

Huggett, M., G. Ventura, and A. Yaron (2011, December). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of political Economy 105*(3), 473–522.

Keane, M. P. and K. I. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Klein, L. and A. Goldberger (1955). An econometric model of the United States, 1929-1952.

Lee, S. Y. and A. Seshadri (2019). On the intergenerational transmission of economic status. *Journal of Political Economy 127*(2), 000–000.

Lochner, L. and A. Monge-Naranjo (2012). Credit constraints in education. *Annu. Rev. Econ. 4*(1), 225–256.

MacKinnon, D. P., A. J. Fairchild, and M. S. Fritz (2007). Mediation analysis. *Annu. Rev. Psychol. 58*, 593–614.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *The Review of Economics and Statistics 87*(2), 235–255.

Meghir, C. and M. Palme (2005). Educational reform, ability, and family background. *American Economic Review 95*(1), 414–424.

Milligan, K. and M. Stabile (2011). Do child tax benefits affect the well-being of children? Evidence from Canadian child benefit expansions. *American Economic Journal: Economic Policy 3*(3), 175–205.

Mogstad, M., A. Fagereng, and M. Ronning (2021). Why do wealthy parents have wealthy children? *Journal of Political Economy*.

Nybom, M. and J. Stuhler (2014). Interpreting trends in intergenerational mobility. Working Paper Series 3/2014, Stockholm University, Swedish Institute for Social Research.

Papageorge, N. W., V. Ronda, and Y. Zheng (2019). The economic value of breaking bad: Misbehavior, schooling and the labor market. Technical report, National Bureau of Economic Research.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, 393–408.

Stinebrickner, R. and T. Stinebrickner (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review 98*(5), 2163–84.

Theil, H. (1958). Economic forecasts and policy.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis.* American Psychological Association.

Todd, P. E. and W. Zhang (2020). A dynamic model of personality, schooling, and occupational choice. *Quantitative Economics 11*(1), 231–275.

A Data

We use only one dataset - the National Child Development Study (NCDS) - to estimate the IGE and the equations for the mediation analysis. The NCDS is a panel tracking a cohort of individuals born in the UK in a particular week in March 1958. In this section, we describe how the parental income and children's lifetime earnings variables are constructed from the NCDS data, as well as how we select our analytical sample.

A.1 Variable Construction

A.1.1 Parental Income

Parental income variables are only observed in the third wave of the NCDS, when cohort members are aged 16. There are three relevant variables: fathers' net earnings, mothers' net earnings and other net income. We drop individuals whose parents did not answer any of the earnings questions. Adopting Blanden et al. (2013)'s procedure, we also drop households where at least one parent is working but that parents' earnings is unobserved. Finally, we restrict age of father at the time of the survey at birth to be between 20 and 45, to ensure that the point-in-time income measure is a reasonable proxy for lifetime income (a detailed breakdown of how we arrive to the final sample can be found in A.2).

Fathers' and mothers' earnings and other net income are recorded as interval, or "banded" data. We use the continuous parental income variables which were imputed by the Institute for Fiscal Studies as part of an effort to harmonize income variables in different cohort studies (Belfield et al. (2017)). This procedure can be summarized by three steps. First, income bands identical to the NCDS are created for the 1973, 1974 and 1975 Family Expenditure Survey (FES) data. Second, within each income band, net male, net female and net other income data are estimated using variables that are thought to be correlated with income.[12]. Third, using these prediction equations, each of the three income components are separately predicted within each income band in the NCDS, using the same key covariates. The sum of the three net, predicted earnings variables in the NCDS is the measure of parental income that we use.

A.1.2 Children's Earnings

NCDS cohort members' gross wages are observed at ages 23, 33, 42, 50, and 55. Gross earnings are calculated using usual wages and pay period on the respondent's main job. Current or last wage variables are used where usual wages are unavailable.

Based on NCDS employment spells data from 1974-2008, we apply three criteria for sample selection as in Gregg et al. (2016). Firstly, cohort members are excluded if less than 60 months of working history is observed. Secondly, cohort members with no earnings observations across the five survey waves are excluded. Finally, cohort members who do not have parental income observation from childhood are not considered in our analysis.

The crucial step in dealing with this sample is in how we interpret missing and zero earnings figures for cohort members, which could be due to attrition, item non-response or non-employment. To address this, we adapt Gregg et al. (2016)'s method of imputing missing earnings at the time of a survey. For this imputation, earnings are modeled as a function of an individual-specific fixed effect and a categorical education variable interacted with age of cohort member:

$$\ln Y_{it} = f_i + \chi \cdot S_i^* \cdot t + \zeta_{it}$$

where $S_i^* = 1$ if less than O level, 2 if O level qualification, 3 if A level qualification, 4 if some college or above.[13] Once earnings at ages 23, 33, 42, 50, and 55 have been predicted as $\widehat{\ln Y_{it}} = \hat{f}_i + \hat{\chi} \cdot S_i^* \cdot t$, the earnings value for individuals recorded as being out of work at the time of a survey is replaced with

[12] The following variables are used to estimate income data within each band: year of interview, age left school of mother and father, employment status of mother and father, age of mother and father, occupation of mother and father, number of siblings of the cohort member, housing tenure, region, number of rooms in household, number of cars, marital status, whether benefits are received by household, type of school cohort member attends, interactions between age left school and occupation of mother and father, interactions between age left school and employment status of mother and father, and interactions between year of interview and occupation of mother and father

[13] O levels and A levels were national exams taken at approximately age 16 and 18 respectively.

zero. This accounts for the fact that missing earnings in those cases represented non-employment, as opposed to attrition. Monthly earnings are calculated from earnings at each survey date by imputing a linear trajectory between each earnings data point. Aggregating monthly earnings provides an estimate of children's lifetime earnings.

A.2 Sample Selection

The initial wave of the NCDS consists of 18,558 individuals. We impose the following requirements for individuals to be in the mediation sample:

- Parental income available (see section A.1.1 for details) - drop 12,170 observations
- Father's age is between 20 and 45 at the time child was born - drop 682 observations
- Children's years of school available - drop 1,572 observations
- Children's lifetime earnings available (see section A.1.2 for details) - drop 117 observations
- Mother and father's years of schooling available - drop 151 observations
- Number of siblings is available - drop 1 observation
- At least one measure of school quality available at each age 7,11,16 - drop 1,048 observations
- At least one measure of time investment available at each age 7,11,16 - drop 106 observations
- At least one measure of ability available at each age 0,7,11,16 - drop 14 observations

The final sample consists of 1350 males and 1347 females.

A.3 Further Descriptives

Table 8 includes descriptive statistics on parental investment measures which we use for our analysis. We also report the gradients by parental income tertile, and F-tests on mean-equivalence across the groups.

Table 8: Investments: Descriptive Statistics and Means by Parental Income Tertile

			Parental Income Tertile			
Variable	Mean	SD	Bottom	Middle	Top	P-values
Time Investments						
Age 7						
% of fathers very interested in child's education	21.5	41.1	18.6	21.5	24.4	0.01
% of mothers very interested in education	34.4	47.5	31.5	34.1	37.4	0.03
% of fathers who go on weekly outings with child	69.8	45.8	65.2	72.5	71.5	0.00
% of mothers who go on weekly outings with child	85.0	35.7	82.8	85.5	86.8	0.07
% of fathers who read to child every week	35.5	47.9	35.9	37.6	32.9	0.12
% of mothers who read to child every week	47.3	49.9	46.0	50.5	45.2	0.06
Age 11						
% of fathers very interested in child's education	23.5	42.4	20.1	23.8	26.4	0.01
% of mothers very interested in child's education	33.4	47.2	29.8	34.3	36.1	0.02
% of fathers who go on weekly outings with child	50.7	50.0	45.0	53.6	53.1	0.00
% of mothers who go on weekly outings with child	53.5	49.9	49.4	56.0	55.1	0.01
% of parents expecting child to finish school	74.5	43.6	70.0	73.6	80.0	0.00
% of parents expecting child to attend university	83.1	37.5	81.2	82.8	85.2	0.08
Age 16						
% of fathers very interested in child's education	29.9	45.8	28.2	28.6	32.9	0.08
% of mothers very interested in child's education	33.3	47.1	31.5	32.8	35.6	0.19
% of parents expecting child to attend university	26.6	44.2	24.1	25.1	30.6	0.01
School Quality Investments						
Age 7						
Private school (in %)	1.45	11.9	1.34	0.89	2.11	0.10
% whose school has a PTA	15.5	36.2	16.9	15.9	13.5	0.11
% whose PTA holds educational meetings	57.7	49.4	56.8	57.6	58.7	0.71
Social class of fathers in child's class	3.04	0.61	3.09	3.07	2.98	0.00
% of children discussed in PTA meetings	48.7	31.1	45.8	48.9	51.4	0.00
Age 11						
Private school (in %)	1.00	10.1	0.92	0.58	1.66	0.08
Class size	35.0	6.70	34.5	35.1	35.5	0.01
Student-teacher ratio	24.8	9.35	25.4	24.7	24.3	0.06
% from child's class suitable for GCEs	23.6	14.0	23.4	23.1	24.3	0.20
Age 16						
Private school (in %)	1.93	13.5	1.22	0.98	3.29	0.00
% from child's school who complete school	55.7	25.8	53.5	54.2	59.6	0.00
% from child's school now in full-time degrees	1.02	1.71	0.91	0.95	1.19	0.00
% from child's school who passed 2+ A-levels	1.84	2.76	1.68	1.72	2.12	0.00
% from child's class studying for GCEs	46.3	31.2	44.0	44.4	50.5	0.00
Student-teacher ratio	17.2	1.93	17.2	17.3	17.1	0.05

Note: This table presents descriptive statistics and means by parental income tertile for all of the investment measures we use. The relevant sample consists of 2,637 individuals. The final column reports *p-values* from the *F-tests* testing the null hypothesis of equality of means across income tertiles.

B Measurement Error in Parental Income

Our goal is to estimate the IGE, which is the slope coefficient in the OLS regression of child's log lifetime earnings on parent's log lifetime income, where both variables are de-meaned. It is ρ in the following regression:

$$\ln Y_i = \rho \ln Y_{Parent,i} + u_i$$

However, parent's lifetime income is unobserved in the NCDS. It is proxied using income observed in the year the children turn 16 (and median age of their parents is close to 42). Previous research has shown that attenuation bias from such proxying can be significant (Haider and Solon (2006)). To address this problem, we use the errors-in-variables framework described below.

First, consider the log-linear projection of age 42 income on lifetime income for parents:

$$\ln Y_{Parent,i,42} = \varphi_{Parent} \ln Y_{Parent,i} + \varepsilon_{Parent,i} \tag{12}$$

Allowing φ_{Parent} to be different than one makes this framework more general than the textbook errors-in-variables framework. We assume the following:

- $Cov(\ln Y_{Parent,i}, u_i) = 0$
- $Cov(\ln Y_{Parent,i}, \varepsilon_{Parent,i}) = 0$
- $Cov(u_i, \varepsilon_{Parent,i}) = 0$

Using this projection and these assumptions, the slope coefficient that we can estimate using observables is ρ^*, the slope coefficient from the regression of child's lifetime earnings on parents' age 42 income:

$$\ln Y_i = \rho^* \ln Y_{Parent,i,42} + u_i$$

Note that by substituting for $\ln Y_i$ using the IGE regression and then substituting for $\ln Y_{Parent,i,42}$ using the log-linear projection we can rewrite ρ^* to be:

$$\rho^* = \frac{\mathbb{E}[\ln Y_i \ln Y_{Parent,42} - \mathbb{E}[\ln Y_{Child}]\mathbb{E}[\ln Y_{Parent,42}]]}{\mathbb{E}[\ln Y_{Parent,42}{}^2 - \mathbb{E}^2[\ln Y_{Parent,42}]]} = \rho \frac{\varphi_{Parent} Var(Y_{Parent})}{\varphi^2_{Parent} Var(Y_{Parent}) + Var(\varepsilon_{Parent})} \tag{13}$$

Thus, ρ^* is equal to the true IGE multiplied by the reliability ratio, defined:

$$RR_{Parent} = \frac{\varphi_{Parent} Var(Y_{Parent})}{\varphi^2_{Parent} Var(Y_{Parent}) + Var(\varepsilon_{Parent})}$$

We can estimate ρ^* using available NCDS data. However, this does not identify ρ because the reliability ratio for the parent generation, RR_{Parent}, is unobserved. Our approach to identify RR_{Parent} is to assume that the relationship between point-in-time income and lifetime income is the same across both generations, that is, $\varphi_{Parent} = \varphi_{Child}$. We also assume that the variance of measurement error in lifetime income when using point in time income as a proxy is the same across the two generations, i.e. $Var(\varepsilon_{Parent}) = Var(\varepsilon_{Child})$. Here, φ_{Child} and $Var(\varepsilon_{Child})$ are identified using equation (12) for the children. We then calculate $Var(Y_{Parent})$ from equation (12) and arrive at an estimate of RR_{Parent}. The final step is to divide our estimate of ρ^* in equation (13) by RR_{Parent} to arrive at ρ. Adjusting our main estimates for measurement error pushes the IGE up from 0.155 to 0.317 for males and from 0.115 to 0.236 for females.

C Different measures of intergenerational elasticity of earnings and income

In our estimates of the intergenerational elasticity of earnings we use children's gross (pre-tax) lifetime earnings and net (post-tax) parental income (father's and mother's net earnings, and other net income) when children are 16. This choice is made because the NCDS only reports net income for parents, while gross earnings for children are found to haver significantly less attrition than net earnings for children. Moreover, the components included in the net income measures are different for parents and children.

In this appendix, we report five alternative estimates of the IGE, and compare them to our preferred measures. None of the measures here are corrected for measurement error.

Table 9 presents our different versions of the IGE. Our main estimates without measurement error corrections for parental income are 0.155 and 0.115 for males and females, respectively. As we discuss in the main text, corrected for measurement error, the estimates become 0.317 and 0.236.

Using a simple mean of point-in-time earnings over the time we see them to construct children's lifetime earnings instead of our fixed-effects imputation pushes our estimate of the IGE down by 3.5 percentage points for males and pushes it up by 3 percentage points for females. This is reported in row (2) of Table 9.

In our main estimate, we predict the three components of parent's income within each income band, using characteristics that are correlated with income. An alternative procedure is to directly calculate median income within each income band from the FES, and replace interval data in the NCDS with the corresponding medians. Row (3) of Table 9 presents the IGE using this alternative measure of parents' income. This reduces our estimate of the IGE by no more than 2 percentage points. Gregg et al. (2016) find an estimate of 0.18 for males when they use this construction of parents' income, although their sample selection criteria differ from ours. Using their criteria, we obtain an IGE estimate of 0.178 when median income is used to construct parental income.

Table 9: Different measures of the IGE

	Measure of children's earnings	Measure of parent's earnings	IGE for Males	IGE for Females
	Full Sample			
(1)	Gross lifetime earnings	Within-band imputed net income	0.155	0.115
			(0.045)	(0.050)
(2)	Gross mean annual earnings	Within-band imputed net income	0.120	0.141
			(0.053)	(0.050)
(3)	Gross lifetime earnings	Within-band median net income	0.139	0.106
			(0.050)	(0.053)
	N		1350	1347
	Restricted Sample			
(4)	Gross lifetime earnings	Within-band imputed net income	0.211	0.047
			(0.063)	(0.089)
(5)	Net lifetime earnings	Within-band imputed net income	0.179	0.090
			(0.043)	(0.059)
(6)	Net family lifetime earnings	Within-band imputed net income	0.096	0.122
			(0.045)	(0.044)
N		404	377	

Note: Row (1) presents our main baseline estimates of the IGE using children's gross lifetime earnings and net parental income when children are 16. Each subsequent row changes one aspect of the estimation, holding all other aspects constant. Gross mean annual earnings is the simple average of children's gross earnings at 23, 33, 42, 50 and 55 multiplied by length of child's working life. Within-band median income uses medians for each band from the FES when imputing for net parental income at 16. Net lifetime earnings is constructed by using take-home pay at 23, 33, 42 and 50 for children and our fixed-effects imputation. Net lifetime family earnings sums across the imputed net lifetime earnings of the child and their partner. Note, row (4) re-estimates the same IGE measure as in row (1) but on a restricted sample, for comparison with estimates in rows (5) and (6).

Belfield et. al (2017) make a distinction between net and gross earnings for children. Our baseline uses gross lifetime earnings, while the estimate in row (5) of Table 9 uses the child's net lifetime earnings. To construct net lifetime earnings, we simply use take-home pay at ages 23, 33, 42 and 50 instead of gross wage. The NCDS measures take-home pay as "pay after deductions for tax and National Insurance, including any overtime, bonus, commission or tips". Since data on take-home pay is only available for some observations in our main sample, we re-estimate our baseline IGE using the restricted sample for comparison (this is presented in row (4)). This gives us 0.211 and 0.047 for males and females, respectively. Using net instead of gross lifetime earnings for children pushes our IGE for males down to 0.179 , and pushes the IGE for females up to 0.090, both relative to our baseline calculated on the restricted sample.

A final way to estimate the IGE is to use family income for the children as well as for the parents. Belfield et. al (2017) find an IGE of 0.203 when they use son's gross family income at 42 and parent's net income from the NCDS. We find estimates of 0.096 for males and 0.122 for females but we use net lifetime family income for children, which is the sum of net lifetime earnings of the NCDS cohort member and their partner.

Overall, we find that the uncorrected IGE is relatively robust to the measure of child's earnings used. We prefer using our fixed effect imputed child gross lifetime earnings measure, as it minimizes problems from attrition, and is the most accurate reflection of lifetime earnings of the child.

D Estimation

D.1 Choice of Measures and Number of Factors - Exploratory Factor Analysis

As described in Section 4.3.1, we conduct an exploratory factor analysis (EFA) to select measures of investments and human capital. EFA is a statistical technique for data reduction. It reduces the number of variables in an analysis by describing linear combinations of the measures using a smaller number of "factors" that contain most of the information. By taking the eigenvalues of the data matrix of measures it allows us to understand how the different measures are related to one another. Based on the Kaiser rule, which suggests to only retain only factors with an eigenvalue greater than one, we retain only two factors which we label time investments and school quality.

Table 10 presents the result for investments. As can be seen from the tables, for ages 11 and 16, the measures clearly load onto two different dimensions: one that is related to parental behaviours, and another that is related to the school environment. It is also noticeable that both dimensions of investments have eigenvalues greater than one. One variable at age 16 – whether parents wish for their child to continue to further education – loads similarly on both time investments and school quality. As we are using a dedicated measurement system where each measure can only load onto one factor, we drop this measure. At age 7, parental interest in education does not load onto the same factor as outings and reading with the child. However, in order to ensure that measures capture the same underlying factor at different ages, we decide to assign it to the time investment category (as in the other periods) rather than the school quality investment category in the final analysis. Classroom size at age 7 loads similarly on both time and school quality, hence we drop it from the final analysis. Finally, following Heckman et al. (2013), we exclude measures with low loadings from the final analysis.

Turning to human capital at age 16, Table 11 present the results of the EFA. We see the measures load onto what is commonly considered a cognitive and a non-cognitive dimension of human capital. It should be noted that only the first two eigenvalues are greater than one, thus indicating that there is no third dimension (such as a second non-cognitive dimension) for this sample. Again, measures with low loadings are excluded from the final analysis.

Table 10: Exploratory Factor Analysis of Investment Measures

Age 7			Age 11			Age 16		
	Factor 1	Factor 2		Factor 1	Factor 2		Factor 1	Factor 2
	SQ	TI		SQ	TI		SQ	TI
Eigenvalues:	1.88	0.93		2.01	1.35		4.27	1.66
Loadings:								
F:interest in ed	0.65	0.01	**F:interest in ed**	0.11	0.62	**F:interest in ed**	0.01	0.87
M:interest in ed	0.68	0.04	**M:interest in ed**	0.09	0.61	**M:interest in ed**	-0.03	0.92
F: outings	0.01	0.63	**F: outings**	-0.15	0.55	**P:support schooling**	0.05	0.60
M:outings	-0.01	0.58	**M:outings**	-0.14	0.53	P:wish ct ed	0.30	0.41
F:reads	-0.02	0.61	**P: wish ct ed**	0.02	0.44	**school type**	0.45	0.04
M:reads	0.03	0.54	**P: wish uni**	0.00	0.35	**%Ct School**	0.63	0.12
school type	0.35	-0.14	**class size**	-0.40	0.10	**%FT degree**	0.89	-0.04
avg soc class of school	-0.47	0.05	**school type**	0.86	-0.01	**%GCE (past)**	0.94	-0.05
PT meetings available	0.19	-0.04	**% GCE**	0.49	0.19	**%GCE (present)**	0.70	0.04
Parental engagement	0.34	-0.06	**teacher-student ratio**	0.61	-0.02	**teacher-student ratio**	0.50	-0.07
PTA	0.17	-0.04	teacher experience	0.14	-0.01	avg soc class of school	0.62	0.13
class size	-0.17	0.12	teacher initiative	0.00	0.25	class size	-0.12	0.25
						career guidance	0.11	-0.03
						hours of english	-0.20	0.04
						hours of maths	-0.10	0.10

Note: The top panel presents eigenvalues of the exploratory factor analysis. The bottom panel presents loadings of the exploratory factor analysis. Measures that are subsequently included in the main analysis are in **bold**. Factor loadings rotated. 'F' denotes father, 'M' denotes mother, 'P' denotes parents.

Table 11: Exploratory Factor Analysis of Skills at 16

	Factor 1	Factor 2
	Cognitive	Non-Cognitive
Eigenvalue:	3.215	1.648
Loadings:		
Reading score	0.78	0.01
Math score	0.81	0.01
Teacher-rated Math ability	0.88	0.00
Teacher-rated English ability	0.87	-0.01
Unconcentrated	0.09	0.46
Destructive	0.03	0.30
Miserable	-0.05	0.52
Fidgety	0.00	0.44
Worried	-0.14	0.44
Irritable	0.01	0.54
Fights	0.06	0.44
Disobedient	0.10	0.45
Solitary	-0.11	0.25
Bullied	-0.02	0.29
Sucks on thumb	-0.02	0.14
Bites nails	0.04	0.21
Upset	-0.05	0.33
Has mannerisms	-0.02	0.21

Table presents loadings for exploratory factor analysis. Measures that are subsequently included in the main analysis are in **bold**.

D.2 Estimating the Cognition, Parental Investment, Schooling Choice, and Lifetime Earnings Equations

Here we discuss how to estimate the production function for: cognition at 16, investment decision rules, schooling choices, and lifetime earnings equations. Estimating these equations is not straightforward because both both the left- and the right-hand-side variables in the equations are latent factors that we only observe through multiple error ridden measures. We closely follow Heckman et al. (2013) and use an errors-in-variables approach to address measurement error.

There are three main steps when using our noisy measures of skills, school quality, investments, and parental income. First, we establish how each individual measure relates to the underlying unobserved latent factor. This is what we lay out in Sections D.2.1, D.2.2 and D.2.3, where we describe the measurement system, the statistical assumptions on the measurement system, and estimation of the measurement system. Second, we combine these different measures into one comprehensive index that can be used for estimation. We do this by estimating Bartlett scores, described in Section D.2.4. Third, in Section D.2.5 we correct for the remaining measurement error in the Bartlett scores using an errors in variables formula.

D.2.1 Measurement System

We do not observe parental income (Y_{Parent}), children's skills (C), parental time investments (inv), or school quality (sq) directly, and thus they are latent factors. Instead, we observe multiple error-ridden measurements of each.[14] These measures have arbitrary scale and location. In particular:

$$Z_{\omega,i,t,j} = \mu_{\omega,t,j} + \lambda_{\omega,t,j}\omega_{i,t} + \epsilon_{\omega,i,t,j} \tag{14}$$

for $\omega \in \{Y_{Parent}, C, inv, sq\}$ and each $j = \{1, ..., J_{\omega,t}\}$ error-ridden measurements of each latent factor ω.

D.2.2 Assumptions on Measurement Errors and Normalizations

Measurement Errors Measurement errors are assumed to be independent across individuals, measures, and time. Measurement errors are also assumed to be independent of the latent variables, and all other controls and shocks. In particular, we make the following assumptions on our measurement model:

1. $\epsilon_{\omega,t,j} \perp \epsilon_{\omega,t,j'}$ for all t, ω and $j \neq j'$
2. $\epsilon_{\omega,t,j} \perp \epsilon_{\omega,t',j'}$ for all ω and $t \neq t'$ and j, j'
3. $\epsilon_{\omega,t,j} \perp \omega'_{t'}$ for all ω, ω', t, t', and j

[14] We do not have multiple measures of parents' income, but we show that we can use multiple measures of children's earnings in our errors in variables framework in appendix B.

4. $\epsilon_{\omega,t,j} \perp X_{t'}$ for all ω, X, t, t'

5. $\epsilon_{\omega,t,j} \perp u_{t'}$ for all ω, t, t' where $u_{t'}$ represents a structural shock (for investment, cognition, schooling)

Although we drop the i subscripts for notational convenience, all of the above independence assumptions hold for each individual i. Furthermore, all measurement errors are assumed independent across individuals.

Assumption 1 is that measurement errors are independent contemporaneously across measures. Assumption 2 is that measurement errors are independent over time. Assumption 3-5 are that measurement errors in any period are independent of the latent factors (Assumption 3), covariates (Assumption 4) and structural shocks (Assumption 5) in any period. While these assumptions are strong, they are common in the literature.

Normalizations Our latent factors do not have a natural scale or location. We thus normalize the latent factors to have mean zero and variance one in every period:

$$Var(\omega_t) = 1 \tag{15}$$

$$E[\omega_t] = 0 \tag{16}$$

D.2.3 Estimation of Measurement Parameters

Using the measurement system in appendix D.2.1 and the statistical assumptions made in appendix D.2.2, here we describe the procedure to estimate the measurement parameters.

1. **Location parameters (μs) in measurement equations**

   In all periods, we normalize the mean of the latent factors C, inv, and sq to zero. Furthermore, we de-mean all of our measures, $Z_{\omega,t,j}$. Therefore, using equation (14), we note that:

   $$\mu_{\omega,t,j} = \mathbb{E}[Z_{\omega,t,j}] = 0, \tag{17}$$

2. **Variance of latent factors**

   Using equation (14) and our normalization of the variance of the latent factor, we can derive the scale parameters of each of the latent factors from the covariances of the observed measures:

   $$Cov(Z_{\omega,t,j}, Z_{\omega,t,j^*}) = \lambda_{\omega,t,j}\lambda_{\omega,t,j^*}\, Var(\omega_t) \tag{18}$$

Note that as long as we have at least three measures, we can identify the scaling parameters. For example, if we have three measures, we have three covariances available, and three λs to estimate.

3. **Variance of the measurement error**

   Finally, we can estimate the measurement error variance for each measure using the observed variance of our measures, our estimated scaling parameters, and the normalization of the latent factor:

$$Var(Z_{\omega,t,j}) = \lambda^2_{\omega,t,j} Var(\omega_t) + Var(\epsilon_{\omega,t,j}) \tag{19}$$

D.2.4 Predicting Children's Skills, Parental Time Investments and School Quality using Bartlett Scores

We have multiple measures of cognition, time investments, and school quality. In this section, which borrows heavily from Heckman et al. (2013), we show how to use the Bartlett score method to take a weighted average of these measurements. In the Bartlett score method, the weights are constructed so that noise from measurement error is minimized. In particular, the procedure is as follows:

Step 1 Estimate measurement system as described above.

Step 2 For each individual, predict Bartlett score (after demeaning): $\omega_{S,i,t} = (\lambda'_{\omega,\mathbf{t}}\mathbf{\Omega^{-1}}\lambda_{\omega,\mathbf{t}})^{\mathbf{-1}}\lambda'_{\omega,\mathbf{t}}\mathbf{\Omega^{-1}Z}_{\omega,\mathbf{i,t}}$ - where all the objects are replaced by their estimated counterparts. Here $\lambda_{\omega,\mathbf{t}}$ is a $J_{\omega,t} \times 1$ vector of scaling parameters $\lambda_{\omega,t,j}$ of all measures j for factor ω, and $\mathbf{\Omega}$ is a $J_{\omega,t} \times J_{\omega,t}$ diagonal matrix with the variances of the measurement errors on the diagonal. Hence this step is equivalent to estimating a weighted regression of the measurement equation for each individual where the coefficient of interest is $\omega_{i,t}$. The weights ensure that noisier measures receive a lower weight.

Note, however, that the Bartlett scores are still contaminated with measurement error. Taking a weighted average reduces but does not eliminate measurement error so the Bartlett scores themselves can thus be seen as a measure of the true latent factor: $\omega_{S,i,t} = \omega_{i,t} + \xi_{\omega,i,t}$. Hence, we follow Heckman et al. (2013), who use the following measurement error correction approach:

D.2.5 Measurement Error Correction

- Say we are interested in the following outcome equation, where Y_{it} can be an arbitrary object of interest (e.g., years of schooling, skills, investments) : $Y_{i,t} = \alpha\omega_{i,t} + \gamma X_{i,t} + u_{i,t}$
- Now we have the predicted score: $\omega_{S,i,t} = \omega_{i,t} + \xi_{\omega,i,t}$

- Because $\xi_{\omega,i,t}$ is a weighted average of $\epsilon_{\omega,i,t,j}$, the assumptions made in appendix D.2.2 implies: $(\boldsymbol{X}_i, \boldsymbol{\omega}_i) \perp \boldsymbol{\xi}_{\omega,i}, \quad E\left(\boldsymbol{\xi}_{\omega,i}\right) = 0, \quad \mathrm{Cov}(\boldsymbol{\xi}_\omega, \boldsymbol{\xi}_\omega) = \boldsymbol{\Sigma}_{\boldsymbol{\xi}_\omega \boldsymbol{\xi}_\omega}$

- Now OLS estimates are inconsistent because $Y_{i,t} = \alpha\omega_{S,i,t} + \gamma X_{i,t} + u_{i,t} - \alpha\xi_{\omega,i,t}$.

- $\operatorname{plim}\begin{pmatrix}\hat{\alpha}\\ \hat{\gamma}\end{pmatrix} = \begin{pmatrix}\mathrm{Cov}\left(\boldsymbol{\omega}_S, \boldsymbol{\omega}_S\right) & \mathrm{Cov}\left(\boldsymbol{\omega}_S, \boldsymbol{X}\right)\\ \mathrm{Cov}\left(\boldsymbol{X}, \boldsymbol{\omega}_S\right) & \mathrm{Cov}(\boldsymbol{X}, \boldsymbol{X})\end{pmatrix}^{-1}\begin{pmatrix}\mathrm{Cov}(\boldsymbol{\omega}, \boldsymbol{\omega}) & \mathrm{Cov}(\boldsymbol{\omega}, \boldsymbol{X})\\ \mathrm{Cov}(\boldsymbol{X}, \boldsymbol{\omega}) & \mathrm{Cov}(\boldsymbol{X}, \boldsymbol{X})\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}\\ \gamma\end{pmatrix}$

- Note to see more clearly: $\operatorname{plim}\begin{pmatrix}\hat{\alpha}\\ \gamma\end{pmatrix} = \underbrace{\begin{pmatrix}\Sigma_{\omega,\omega} + \Sigma_{\xi_\omega,\xi_\omega} & \Sigma_{\omega,X}\\ \Sigma_{X,\omega} & \Sigma_{X,X}\end{pmatrix}^{-1}\begin{pmatrix}\Sigma_{\omega,\omega} & \Sigma_{\omega,X}\\ \Sigma_{X,\omega} & \Sigma_{X,X}\end{pmatrix}}_{A}\begin{pmatrix}\alpha\\ \gamma\end{pmatrix}$

- As we know the variance of the latent factors, we thus know $\xi_{\omega,i,t}$ and can correct for it. We can then pre-multiply the estimated coefficients by the inverse of the attenuation factor (A^{-1}) to get the consistent coefficients

- We use the approach described above to estimate the schooling equation, the cognition equations, and also the equations for parental investments in time and school quality. We also use it to correct for measurement error in the regression of mediators on parental income.

D.3 Signal-to-Noise Ratios

Table 12 presents the signal-to-noise ratios for the variables included in our final analysis. The signal-to-noise ratio is defined as $SN_{\omega,t,j} = \frac{\lambda^2_{\omega,t,j} Var(\omega_t)}{Var(Z_{\omega,t,j})}$. Intuitively, this is the appropriately scaled variance of the latent factor (signal) to the variance of the measure (signal+noise) and thus describes the information content of each measure. Note that we have normalized the variance of the latent factors as well as of the measures to one. Hence, the factor loadings $\lambda_{\omega,t,j}$ are the square-root of the signal-to-noise ratio.

E Testing Restrictions in the Presence of Measurement Error

To test for the joint significance of groups of coefficients in Section 4.2, we use a measurement-error-corrected *F-test* procedure. We illustrate this procedure here using a simple example. Suppose the model we are interested in is:

$$y = \alpha_1 x_1^* + \alpha_2 x_2^* + u \tag{20}$$

Under the null hypothesis $H_0 : \alpha_2 = 0$:

$$y = \alpha_1^R x_1^* + u_R$$

Table 12: Signal-to-Noise Ratio for Measures Used

Cognition at 16					
Reading	0.56				
Math	0.62				
Teacher assessed math	0.80				
Teacher assessed english	0.72				
Parental Investments					
Age 7		*Age 11*		*Age 16*	
M:Outings	0.29	M:Outings	0.53	F: Interest in ed	0.75
F:Outings	0.37	F:Outings	0.57	M:Interest in ed	0.90
M:Reads	0.31	F:Interest in Ed	0.12	P:Support schooling	0.32
F:Reads	0.29	M: Interest in Ed	0.14		
M:Interest in Ed	0.15	P:Wish cnt ed	0.09		
F:Interest in Ed	0.11	P:Wish uni	0.06		
School Quality					
Age 7		*Age 11*		*Age 16*	
Avg soc class of school	0.06	Class Size	0.10	School Type	0.08
PT meetings available	0.42	School Type	0.42	%Cnt School	0.35
PTA	0.28	% GCE	0.20	%FT degree	0.82
School type	0.01	Teacher student ratio	0.21	% passed A-levels	0.93
Parental engagement	0.10			% studying towards A-levels	0.45
				Teacher student Ratio	0.20

Note: 'F' denotes father, 'M' denotes mother, 'P' denotes parents.

Under the alternative hypothesis $H_1 : \alpha_2 \neq 0$:

$$y = \alpha_1^{UR} x_1^* + \alpha_2^{UR} x_2^* + u_{UR}$$

The standard *F-statistic* is then constructed as:

$$F = \frac{(SSR_R - SSR_{UR})/q}{SSR_{UR}/(n-k-1)}, \tag{21}$$

where $SSR_R = \sum_{i=1}^n (y_i - \hat{\alpha}_1^R x_{1,i}^*)^2$, $SSR_{UR} = \sum_{i=1}^n (y_i - \hat{\alpha}_1^{UR} x_{1,i}^* - \hat{\alpha}_2^{UR} x_{2,i}^*)^2$, $q = 1$ is the number of restrictions, n is the sample size and $k = 1$ is the number of independent variables in the unrestricted model. However, in our model, we do not observe the true values of independent variables (i.e. x_1^* and x_2^*). Instead our observed variables, x_1 and x_2, are measured with an additive error:

$$\begin{aligned} x_1 &= x_1^* + \epsilon_1 \\ x_2 &= x_2^* + \epsilon_2 \end{aligned} \tag{22}$$

Using this framework, the aim is to now express the *F-statistic* in terms of observables only. First, we re-write both the sum of squared residual terms in terms of variances:

$$SSR_R = \sum_{i=1}^{n}(y_i - \hat{\alpha}_1^R x_{1,i}^*)^2 = n \cdot Var(y - \hat{\alpha}_1^R x_1^*) \tag{23}$$

$$SSR_{UR} = \sum_{i=1}^{n}(y_i - \hat{\alpha}_1^{UR} x_{1,i}^* - \hat{\alpha}_2^{UR} x_{2,i}^*)^2 = n \cdot Var(y - \hat{\alpha}_1^{UR} x_1^* - \hat{\alpha}_2^{UR} x_2^*) \tag{24}$$

Substituting the expressions in (22) into the RHS of (23) and (24):

$$\begin{aligned} n \cdot Var(y - \hat{\alpha}_1^R x_1^*) &= n \cdot Var(y - \hat{\alpha}_1^R (x_1 - \epsilon_1)) \\ &= n[Var(y - \hat{\alpha}_1^R x_1) + (\hat{\alpha}_1^R)^2 Var(\epsilon_1) + 2Cov(y - \hat{\alpha}_1^R x_1, \epsilon_1)] \\ &= n[Var(y - \hat{\alpha}_1^R x_1) + (\hat{\alpha}_1^R)^2 Var(\epsilon_1) - 2(\hat{\alpha}_1^R)^2 Var(\epsilon_1)] \\ &= n[Var(y - \hat{\alpha}_1^R x_1) - (\hat{\alpha}_1^R)^2 Var(\epsilon_1)], \end{aligned} \tag{25}$$

$$\begin{aligned} & n \cdot Var(y - \hat{\alpha}_1^{UR} x_1^* - \hat{\alpha}_2^{UR} x_2^*) \\ &= n[Var(y - \hat{\alpha}_1^{UR}(x_1 - \epsilon_1) - \hat{\alpha}_2^{UR}(x_2 - \epsilon_2))] \\ &= n[Var((y - \hat{\alpha}_1^{UR} x_1 - \hat{\alpha}_2^{UR} x_2) + (\hat{\alpha}_1^{UR}\epsilon_1 + \hat{\alpha}_2^{UR}\epsilon_2))] \\ &= n[Var(y - \hat{\alpha}_1^{UR} x_1 - \hat{\alpha}_2^{UR} x_2) + Var(\hat{\alpha}_1^{UR}\epsilon_1 + \hat{\alpha}_2^{UR}\epsilon_2) + 2Cov(y - \hat{\alpha}_1^{UR} x_1 - \hat{\alpha}_2^{UR} x_2, \hat{\alpha}_1^{UR}\epsilon_1 + \hat{\alpha}_2^{UR}\epsilon_2)] \\ &= n[Var(y - \hat{\alpha}_1^{UR} x_1 - \hat{\alpha}_2^{UR} x_2) + (\hat{\alpha}_1^{UR})^2 Var(\epsilon_1) + (\hat{\alpha}_2^{UR})^2 Var(\epsilon_2) - 2(\hat{\alpha}_1^{UR})^2 Var(\epsilon_1) - 2(\hat{\alpha}_2^{UR})^2 Var(\epsilon_2)] \\ &= n[Var(y - \hat{\alpha}_1^{UR} x_1 - \hat{\alpha}_2^{UR} x_2) - (\hat{\alpha}_1^{UR})^2 Var(\epsilon_1) - (\hat{\alpha}_2^{UR})^2 Var(\epsilon_2)], \end{aligned} \tag{26}$$

where we assume the two measurement errors to be uncorrelated, $Cov(\epsilon_1, \epsilon_2) = 0$, and the unobserved factors to be uncorrelated with the measurement errors $Cov(\epsilon_j, x_k^*) = 0$ for all $j = \{1, 2\}$ and $k = \{1, 2\}$.

Expressions (25) and (26) can now be calculated using observed independent variables and estimated measurement error variances. Using (23), (24), (25) and (26), the $F-statistic$ can be expressed in terms of observables.

Within this framework, our procedure to test restrictions in the presence of measurement error is:

1. Estimate α_1 and α_2. Since these are the coefficients from the true model, we estimate them using an errors-in-variables correction. Call these estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$.

2. Obtain residuals $y - \hat{\alpha}_1 x_1$ and $y - \hat{\alpha}_1 x_1 - \hat{\alpha}_2 x_2$ and estimate their variances.

3. Adjust $Var(y - \hat{\alpha}_1 x_1)$ and $Var(y - \hat{\alpha}_1 x_1 - \hat{\alpha}_2 x_2)$ for (estimated) measurement error variances $\hat{Var}(\epsilon_1)$ and $\hat{Var}(\epsilon_2)$ as in (25) and (26). This gives us SSR_R and SSR_{UR}.

4. Calculate the *F-statistic* as in (21) and using estimates of SSR_R and SSR_{UR}.

5. Bootstrap for the distribution of the *F-statistic*. We bootstrap the distribution of the statistic because measurement error make the distribution of the statistic complex.

F Effect of Parental Income on Mediating Variables

Table 13 shows the results of a regression of each mediating variable on log parental income, as illustrated in equation 5. We do this separately for males and females and correct for measurement error in parental income using the correction approach outlined above.

Table 13: Coefficient on Log Parental Income

Dependent Variable	Males	Females
Years of Schooling	0.530	**0.748**
	(0.301)	(0.242)
Cognition	**0.743**	**0.499**
	(0.175)	(0.142)
Time Investments		
Age 7	**0.966**	**0.542**
	(0.185)	(0.204)
Age 11	**1.175**	**0.562**
	(0.202)	(0.188)
Age 16	**0.557**	**0.351**
	(0.165)	(0.155)
School Quality Investments		
Age 7	0.017	0.238
	(0.172)	(0.204)
Age 11	**0.600**	**0.180**
	(0.231)	(0.030)
Age 16	**0.427**	**0.563**
	(0.150)	(0.162)
Family Background		
Mother's education	**0.636**	**1.049**
	(0.159)	(0.190)
Father's education	**0.776**	**1.086**
	(0.218)	(0.238)
Number of siblings	-0.400	-0.176
	(0.235)	(0.245)
N	1350	1347

Standard errors bootstrapped with 250 repetitions.

Note: This table presents estimates of Equation (5) for each determinant we use for lifetime earnings. Each row contains the coefficient (and corresponding standard error) on log parental income from a measurement error corrected regression of the variable in the first column on log parental income, for males and females.

G Results for Robustness Checks

G.1 Robustness Check 1: More Covariates

Table 14 shows results for the mediation analysis when we additionally control for marital stability and parental age. Marital stability is a dummy that takes on value 1 when the child lives with both natural parents from birth to age 16. We find that none of the additional covariates explain a significant fraction of the IGE.

Note, despite the effect of birth order on lifetime outcomes being widely studied, we do not include it here as a mediator for the IGE. The reason for this is that conditional on the number of children, birth order is uncorrelated with parental income. Hence it cannot mediate the effect of parental income on child earnings.

G.2 Robustness Check 2: Including Non-Cognitive Skills

Table 15 is analogous to the mediation analysis Table 4 in the main text, but also includes non-cognitive skills. Non-cognitive skills enter in the same way as cognition; they can affect years of schooling, and are affected by investments and family background. We can see that non-cognitive skills explain a small and insignificant fraction of the IGE. Table 16 shows estimated determinants of lifetime earnings and years of schooling with non-cognitive skills. For both lifetime earnings and years of schooling, cognitive skills matter much more than non-cognitive skills although both are positively related to parental income. Finally, Table 17 shows the relationship between parental income and non-cognitive skills at age 16. We find a much stronger relationship between parental income and child cognition, than between parental income and non-cognitive skills. In fact, for men, non-cognitive skills are negatively correlated with parental income.

Table 14: Main Mediation Analysis Including Additional Covariates

	Males				Females			
	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4
Years of Schooling	**0.095**	-0.096	-0.096	-0.096	**0.423**	0.093	0.093	0.093
	[0.029, 0.203]	[-0.260, -0.001]	[-0.260, -0.001]	[-0.260, -0.001]	[0.141, 1.125]	[-0.280, 0.733]	[-0.280, 0.733]	[-0.280, 0.733]
Cognition	**0.323**	**0.454**	0.149	0.149	0.129	**0.396**	-0.021	-0.021
	[0.168, 0.631]	[0.255, 0.837]	[-0.086, 0.406]	[-0.086, 0.406]	[-0.030, 0.435]	[0.152, 1.142]	[-0.649, 0.319]	[-0.649, 0.319]
Investments	0.134	0.187	**0.469**	0.306	0.049	0.150	**0.449**	0.277
	[-0.104, 0.433]	[-0.043, 0.518]	[0.178, 0.962]	[-0.007, 0.799]	[-0.398, 0.458]	[-0.224, 0.699]	[0.106, 1.579]	[-0.341, 1.202]
Time Investments	0.132	0.178	**0.388**	**0.281**	-0.093	-0.038	0.116	-0.070
	[-0.075, 0.445]	[-0.040, 0.519]	[0.156, 0.826]	[0.055, 0.753]	[-0.427, 0.122]	[-0.321, 0.203]	[-0.088, 0.454]	[-0.467, 0.239]
Age 7	0.135	0.156	0.152	0.084	0.149	0.167	0.181	-0.018
	[-0.055, 0.420]	[-0.046, 0.463]	[-0.054, 0.412]	[-0.038, 0.283]	[-0.140, 0.619]	[-0.053, 0.722]	[-0.062, 0.769]	[-0.366, 0.247]
Age 11	-0.057	-0.030	0.075	0.066	-0.180	-0.176	-0.133	-0.053
	[-0.331, 0.238]	[-0.293, 0.267]	[-0.163, 0.421]	[-0.145, 0.379]	[-0.637, -0.037]	[-0.634, 0.028]	[-0.518, 0.123]	[-0.371, 0.194]
Age 16	0.054	0.052	**0.162**	**0.131**	-0.062	-0.028	0.067	0.002
	[-0.008, 0.148]	[-0.009, 0.144]	[0.071, 0.365]	[0.032, 0.331]	[-0.276, 0.032]	[-0.184, 0.086]	[-0.016, 0.231]	[-0.106, 0.116]
School Quality	0.002	0.010	0.081	0.024	**0.142**	**0.188**	**0.333**	**0.347**
	[-0.124, 0.135]	[-0.121, 0.142]	[-0.076, 0.234]	[-0.106, 0.126]	[0.008, 0.483]	[0.038, 0.698]	[0.139, 1.187]	[0.014, 0.619]
Age 7	-0.001	-0.001	0.000	0.001	0.047	0.044	0.047	0.062
	[-0.022, 0.032]	[-0.021, 0.028]	[-0.025, 0.029]	[-0.035, 0.046]	[-0.060, 0.198]	[-0.050, 0.184]	[-0.054, 0.186]	[-0.092, 0.300]
Age 11	-0.030	-0.028	-0.051	-0.056	0.017	0.019	0.012	0.005
	[-0.158, 0.050]	[-0.161, 0.053]	[-0.205, 0.029]	[-0.252, 0.031]	[-0.069, 0.129]	[-0.101, 0.113]	[-0.182, 0.086]	[-0.143, 0.089]
Age 16	0.033	0.038	**0.132**	0.080	0.078	0.125	**0.274**	**0.280**
	[-0.010, 0.106]	[-0.006, 0.118]	[0.046, 0.265]	[-0.018, 0.215]	[-0.052, 0.349]	[-0.009, 0.456]	[0.099, 0.884]	[0.064, 0.919]
Family Background	-0.205	-0.197	-0.174	-0.011	-0.264	-0.302	-0.183	-0.012
	[-0.131, 0.089]	[-0.146, 0.078]	[-0.089, 0.166]	[0.063, 0.453]	[-0.288, 0.221]	[-0.220, 0.323]	[-0.089, 0.554]	[0.136, 1.174]
Mother's education	-0.045	-0.044	-0.021	0.031	-0.027	-0.009	0.043	0.147
	[-0.154, 0.010]	[-0.162, 0.016]	[-0.126, 0.038]	[-0.054, 0.103]	[-0.341, 0.267]	[-0.258, 0.284]	[-0.228, 0.368]	[-0.072, 0.631]
Father's education	0.012	0.004	0.032	0.084	0.055	0.066	0.113	0.209
	[-0.073, 0.101]	[-0.092, 0.086]	[-0.053, 0.131]	[0.009, 0.194]	[-0.133, 0.297]	[-0.124, 0.293]	[-0.052, 0.442]	[0.041, 0.763]
Number of Siblings	0.013	0.012	0.028	0.077	-0.020	-0.021	-0.011	0.004
	[-0.418, 0.096]	[-0.381, 0.114]	[-0.426, 0.110]	[-0.390, 0.182]	[-0.118, 0.038]	[-0.114, 0.040]	[-0.070, 0.029]	[-0.037, 0.056]
Stable	-0.145	-0.133	-0.150	-0.103	-0.189	-0.243	-0.178	-0.155
	[-0.418, 0.096]	[-0.381, 0.114]	[-0.426, 0.110]	[-0.390, 0.182]	[-0.923, 0.598]	[-0.893, 0.468]	[-0.780, 0.609]	[-0.834, 0.527]
Mum's age	-0.032	-0.028	-0.038	-0.037	-0.099	-0.087	-0.155	-0.213
	[-0.179, 0.089]	[-0.174, 0.090]	[-0.171, 0.092]	[-0.186, 0.100]	[-0.474, 0.105]	[-0.417, 0.100]	[-0.601, 0.039]	[-0.778, 0.025]
Dad's age	-0.008	-0.007	-0.025	-0.063	0.017	-0.008	0.004	-0.004
	[-0.188, 0.149]	[-0.191, 0.147]	[-0.212, 0.121]	[-0.263, 0.076]	[-0.264, 0.342]	[-0.294, 0.222]	[-0.285, 0.281]	[-0.283, 0.232]
N	1350	1350	1350	1350	1347	1347	1347	1347

Note: This replicates the mediation analysis in Table 4, except it includes additional family background variables.

Table 15: Main Decomposition with Non-Cognitive Skills

	Males				**Females**			
	Level 1	Level 2	Level 3	Level 4	Level 1	Level 2	Level 3	Level 4
Years of Schooling	**0.104**	**-0.078**	**-0.078**	**-0.078**	**0.420**	0.039	0.039	0.039
	[0.031, 0.266]	[-0.274, -0.012]	[-0.274, -0.012]	[-0.274, -0.012]	[0.194, 1.127]	[-0.171, 0.329]	[-0.171, 0.329]	[-0.171, 0.329]
Cognition	**0.338**	**0.474**	0.107	0.107	0.135	**0.394**	0.012	0.012
	[0.181, 0.759]	[0.296, 1.007]	[-0.096, 0.378]	[-0.096, 0.378]	[-0.016, 0.400]	[0.161, 1.071]	[-0.297, 0.212]	[-0.297, 0.212]
Non-cognitive skills	-0.004	-0.005	- 0.046	- 0.046	0.000	0.000	- 0.022	-0.022
	[-0.079, 0.042]	[-0.082, 0.043]	[-0.169, 0.007]	[-0.169, 0.007]	[-0.047, 0.039]	[-0.073, 0.061]	[-0.151, 0.022]	[-0.151, 0.022]
Investments	0.123	0.178	**0.517**	**0.354**	0.033	0.128	**0.444**	0.239
	[-0.133, 0.454]	[-0.063, 0.623]	[0.212, 1.346]	[0.112, 0.974]	[-0.306, 0.351]	[-0.142, 0.525]	[0.158, 1.278]	[-0.038, 0.745]
Family Background	-0.008	-0.018	0.051	0.214	-0.006	0.020	0.108	0.314
	[-0.173, 0.109]	[-0.194, 0.103]	[-0.093, 0.188]	[0.092, 0.558]	[-0.238, 0.215]	[-0.189, 0.303]	[-0.066, 0.504]	[0.089, 0.997]
N	1339	1339	1339	1339	1336	1336	1336	1336

Note: This replicates the mediation analysis in Table 4, except it includes Non-cognitive Skills.

Table 16: Determinants of lifetime earnings and years of schooling with non-cognitive measures

	Lifetime Earnings		**Years of Schooling**	
	Males	Females	Males	Females
Years of Schooling	0.055	0.133		
	(0.014)	(0.015)		
Cognition	0.135	0.064	0.947	0.927
	(0.028)	(0.039)	(0.057)	(0.066)
Non-Cognitive Skills	0.038	0.013	0.076	0.099
	(0.028)	(0.024)	(0.051)	(0.049)
Time Investments				
Age 7	0.036	0.066	0.110	0.038
	(0.036)	(0.044)	(0.085)	(0.081)
Age 11	-0.010	-0.078	0.111	0.013
	(0.035)	(0.043)	(0.101)	(0.080)
Age 16	0.033	-0.045	-0.018	0.164
	(0.024)	(0.026)	(0.056)	(0.050)
School Quality Investments				
Age 7	-0.007	0.048	0.013	-0.016
	(0.023)	(0.026)	(0.051)	(0.049)
Age 11	-0.013	0.023	0.009	0.021
	(0.033)	(0.032)	(0.058)	(0.255)
Age 16	0.024	0.037	0.070	0.158
	(0.020)	(0.023)	(0.048)	(0.066)
Family Background				
Mother's age left school	-0.024	-0.008	0.016	0.037
	(0.018)	(0.023)	(0.040)	(0.055)
Father's age left school	0.011	0.016	-0.067	0.013
	(0.019)	(0.019)	(0.037)	(0.036)
Number of Siblings	-0.013	0.027	0.030	0.001
	(0.013)	(0.016)	(0.031)	(0.032)
Log parental income	0.059	0.047	-0.183	0.033
	(0.048)	(0.048)	(0.103)	(0.107)
N	1339	1336	1339	1336

Note: This table reports estimates from the same model presented in Table 5, except here variables of non-cognitive skills, internalising behaviour and externalising behaviour are included. These are shown to have negligible effect on log lifetime earnings once cognition at 16 and years of schooling are controlled for.

Table 17: Effect of log parental income on cognition and non-cognitive skills

	Males		Females	
	Cognition	Non-Cognitive Skills	Cognition	Non-Cognitive Skills
Log parental income	0.743	-0.032	0.499	0.001
	(0.168)	(0.181)	(0.159)	(0.191)
N	1339	1339	1336	1336

Note: Estimates of the direct effects of log parental income on cognition and non-cognitive skills for males and females at age 16.

G.3 Robustness Check 3: Including Interactions

The errors-in-variables correction we are using is suitable only for linear models. In order to conduct a robustness check which allows for an interaction between cognition and schooling in the lifetime earnings equation, we estimate the lifetime earnings equation using a system GMM approach where we use the different measures as instruments for each other to correct for measurement error following Bolt et al. (2020). This approach extends the one in Agostinelli and Wiswall (2016). The downside of this approach is that we cannot correct for measurement error in parental income anymore (as we only have one measure of parental income). Thus, we first estimate a version of the level 1 mediation where we do not correct for measurement error in parental income. Other than that, the approach is the same as in the main text. Column 1 of Table 18 shows that Level 1 mediation results remain largely unchanged compared to the baseline model, even if we do not correct for measurement error in parental income. Column 2 of Table 18 then provides a comparison between the GMM approach and errors-in-variables, when we assume only linear effects. We can see that the results of using GMM to correct for measurement error or using errors-in-variables to correct for measurement error yields very similar results. Finally, in column 3, we add the interaction between years of schooling and cognition. The interaction term explains a small and insignificant fraction of the IGE, both for males and females. Although we find evidence of an interaction between years of schooling and cognition in the earnings equation, this interaction is not highly correlated with parental income.

Table 18: Robustness including interaction terms

	Males			Females		
	EIV	GMM	GMM	EIV	GMM	GMM
Years of Schooling	**0.093**	**0.165**	**0.162**	**0.425**	**0.452**	**0.487**
	[0.019, 0.228]	[0.073, 0.325]	[0.066, 0.310]	[0.158, 1.337]	[0.231, 1.083]	[0.265, 1.206]
Cognition	**0.333**	**0.368**	**0.365**	0.135	0.094	0.078
	[0.193, 0.729]	[0.173, 0.646]	[0.184, 0.625]	[-0.008, 0.502]	[-0.058, 0.268]	[-0.081, 0.229]
Years of Schooling × Cognition			-0.016			0.003
			[-0.066, 0.017]			[-0.054, 0.070]
Investments	0.163	0.137	0.122	0.057	0.149	0.122
	[-0.060, 0.456]	[-0.112, 0.428]	[-0.119, 0.392]	[-0.266, 0.437]	[-0.140, 0.554]	[-0.124, 0.513]
Family Background	-0.012	-0.055	-0.053	0.022	0.055	0.102
	[-0.150, 0.112]	[-0.232, 0.074]	[-0.215, 0.077]	[-0.233, 0.302]	[-0.164, 0.297]	[-0.136, 0.374]

H Further Comparisons with the Literature

Here we present some further comparisons with estimates in the literature.

School Quality: Altonji and Dunn (1996) use within-family variation in where siblings attended schools, instrumented using family residence, to estimate the direct return to school quality. Interestingly, they find that if anything, OLS estimates understate the return to school quality. In their IV estimates they find that a one standard deviation improvement in the quality of the child's school increases earnings by approximately 10.5%. [15]This is bigger than, but not statistically different from, their OLS estimate. Using the NCDS data, we estimate an identical model to the OLS regression used by Altonji and Dunn (1996) (see Column 1 of Table 19). We find that one standard deviation increase in school quality at all ages increases lifetime earnings by 8.9%. Thus, our estimated impact of school quality is in a similar range as Altonji and Dunn (1996)'s preferred IV estimates.

Family Size: Several papers that use twins as a source of plausibly exogenous variation tend to find small effects of family size on earnings, and suggest that OLS estimates overstate the effect of family size. For example, Black et al. (2005) use Norwegian administrative data on twins and find that an increase in the number of siblings leads to a change in earnings of -2% using OLS, but of 0% for males and 3% for females when using twins as an instrument. [16] These IV estimates are not statistically significant. In Table 19, we use the same controls as Black et al. (2005) and find a larger effect of family size on earnings of -3.1% (see Column 2 of Table 19), implying that we may potentially be over-stating the effect of family size. This observation is also supported by Angrist et al. (2010), who show that the effects of family size on earnings are in the range -2% to +2% when using twins and/or sex composition as instruments, and are -3% when using OLS [17]. However, Daruich and Kozlowski (2020) find that an equalisation of fertility across the income gradient would increase income persistence by 6% - an estimate very close to ours.

[15]See Table 1 in Altonji and Dunn (1996). To construct this estimate, we use the level coefficient from Column 1a of their paper - 0.501 (1.17). Since their coefficients are scaled by 100, we divide by 100, and then multiply by 2.1 (which is SD of the school-quality index).

[16]See Table 9 in the appendix of Black et al. (2005). These estimates are taken from the log(earnings). Row 1 contains relevant OLS estimates, and row 3 contains IV estimates.

[17]OLS estimates are taken from Tables 5 and 6 of Angrist et al. (2010). Pooled IV estimates are from Table 7. The relevant dependent variable in Angrist et al. (2010) is log earnings.

Table 19: Estimated effects of school quality on lifetime earnings

	Children's Lifetime Earnings	
	(1)	(2)
School Quality at 7	0.023	
	(0.011)	
School Quality at 11	0.013	
	(0.011)	
School Quality at 16	0.053	
	(0.015)	
Number of siblings		-0.031
		(0.011)
Constant	0.000	12.999
	(0.014)	(0.258)

Column (1) presents OLS estimates from a regression of children's lifetime earnings on school quality at all ages. It includes the following controls: a cubic in years of schooling, years of schooling interacted with parental education, gender, ethnicity and region of residence during childhood. This replicates the controls used in Altonji and Dunn (1996) as relevant for our data. Column (2) has estimates from OLS regression of children's lifetime earnings on number of siblings, controlling for: parents' age at birth, parental education, gender and birth-order. This replicates the controls used in Black et al. (2005) as relevant for our data.