Intergenerational Altruism and Transfers of Time and Money: A Lifecycle Perspective*

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

February 15, 2019

PRELIMINARY AND INCOMPLETE - PLEASE DO NOT CITE

Abstract

Parental investments in children can take one of three broad forms: (1) Time investments during childhood and adolescence that aid child development, and in particular cognitive ability (2) Educational investments that improve school quality and hence educational outcomes (3) Cash investments in the form of inter-vivos transfers and bequests. We develop a dynastic model of household decision making with intergenerational altruism that nests a child production function, incorporates all three of these types of investments, and allows us to quantify their relative importance and estimate the strength of intergenerational altruism. Using British cohort data that follows individuals from birth to retirement, we find that around 40% of differences in average lifetime income by paternal education are explained by ability at age 7, around 40% by subsequent divergence in ability and different educational outcomes, and around 20% by inter-vivos transfers and bequests received so far.

*Preliminary, please do not quote. For excellent research assistance we thank Jack Light and for helpful comments we thank Andrew Hood, George Levi-Gayle, Limor Golan, and Ananth Seshadri. 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) and Grant Inequality and the insurance value of transfers across the life cycle (ES/P001831/1)) for this work is gratefully acknowledged. O'Dea additionally acknowledges funding from the ESRC Secondary Data Analysis grant ref: ES/N011872/. Correspondence to eric.french.econ@gmail.com or cormac.odea@yale.edu. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium. The opinions and conclusions expressed are solely those of the author and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof. Any errors are the authors' own.

†University College London and Institute for Fiscal Studies

‡University College London and Institute for Fiscal Studies

§University College London

¶Yale University and Institute for Fiscal Studies

1 Introduction

Intergenerational links are a key determinant of levels of inequality and social mobility, with previous work looking at a range of developed economies finding very significant intergenerational correlations in education, incomes and wealth (e.g. Dearden et al. (1997), Mazumder (2005), Charles and Hurst (2003), Chetty et al. (2014)). The literature on understanding the mechanisms behind this persistence is much newer. Understanding the drivers of this persistence of economic outcomes across generations is crucial for the design of tax and transfer policies for two main reasons. First, insofar as the correlations reflect differential parental investments in children (both of time and money) they represent an important reason that the design of public policy should not treat the distributions of ability, education, earnings and wealth as fixed. Policies designed to mitigate the intergenerational transmission of inequality through one channel (e.g. estate taxes) could, by affecting parental investments, increase transmission through another channel (e.g. parental spending on children's education). Second, the extent of parental investment in children over the course of their lives provides important evidence on the extent of intergenerational altruism - the extent to which parents forgo consumption and leisure to invest in their children allows us to estimate the relative weight they put on their children's welfare relative to their own. The degree of intergenerational altruism is a key parameter for assessing the potential benefits of social security and tax reform, since current generations will only be willing to accept cuts to their benefits in order to reduce budget deficits if they are altruistic towards future generations (Fuster et al. (2007)).

In this paper we develop a dynastic model of household decision making that incorporates three different types of parental investment in children: i) time investments during childhood and adolescence that aid child development, and in particular cognitive ability, ii) educational investments that improve school quality and hence educational outcomes and iii) cash investments in the form of inter-vivos transfers and bequests. The key contribution of the paper is to estimate such a model using unique longitudinal data from a survey that has been running for 60 years – following a cohort of individuals from birth to retirement. Using these data, we can measure parental inputs over the whole life cycle, and hence look directly at early life investments of time and goods, estimate a child production function for cognitive ability and link that ability measure to earnings in adulthood. The data also include detailed information about the schooling received by individuals and the inter-vivos transfers they then receive from parents during early adult life.

Using these data, we are able to build and estimate a model capable of speaking to the issues raised above. First, we can provide an estimate of the degree of intergenerational altruism drawing on data on a number of different investment decisions. Such an estimate is likely to be more robust than one based on a single decision (such as how much to leave in bequests) which is likely to be affected by a number of other confounding factors. Second, having estimated the degree of intergenerational altruism (along with the other structural parameters that govern household behaviour), we can run policy counterfactuals and look at how each type of parental investment would respond.

Preliminary analysis of this cohort data suggests that around 40% of differences in average lifetime income by paternal education are explained by ability at age 7, around 40% by subsequent divergence in ability and different educational outcomes, and around 20% by inter-vivos transfers and bequests received so far. These findings are supported by results from a preliminary simple version of the model that has been calibrated to match wealth and labour supply moments. Using consumption equivalent variation to measure the welfare gains from higher-educated parents, we again find that differences in investments before and after the age of 7 are of roughly equal importance in determining lifetime utility differences between children of high versus low educated parents, with investments in ability and education looking much more important than differences in the level of inter-vivos transfers and bequests. Looking in more detail at investments in ability, we find that higher levels of time investments increase ability, and that the ability production function looks to exhibit dynamic complementarity, at least at younger ages (see Cunha et al. (2010)).

Finally, we present estimates of many of the investments that households make in their children, including time and money investments. We show that increased investment of time and goods of parents leads to higher ability children (as measured by test scores), and this higher ability leads to higher wages and incomes later in life. Furthermore, we show that higher income parents invest more in their children, and that these investments can explain much of the difference in lifetime incomes of children across the parental education distribution.

This paper relates to a number of different strands of the existing literature, including work measuring the drivers of inequality and intergenerational correlations in economic outcomes, the large literature seeking to understand child production functions and work on parental altruism and bequest motives. The most closely related papers, however, are those focused on the costs of and returns to parental investments in children. Our paper is most similar to Lee and Seshadri (2016). They develop a model that has many similar features to that used in our paper, but they lack data that links investments at young ages to earnings at older ages. As a result, they have to calibrate key parts of the model, while we are able to estimate the human capital production technology, and show how early life investments and the resulting human capital impacts late life earnings. Caucutt and Lochner (2012) is also related to ours. Their paper estimates a human capital production function and altruistic parental transfers to improve human capital of children. They find that borrowing constraints are an important deterrent to college going. They use data on parental investments at different ages and also later life income, but, unlike us, they they cannot directly measure early life ability however. Furthermore, they restrict the set of investments that can be made in children because they do not allow for endogenous labor supply or inter-vivos cash transfers. Other closely related papers include Del Boca et al. (2014) and Gayle et al. (2015), both of which develop models in which parents choose how much time to allocate to the labour market, leisure and investment in children. Neither paper, however, incorporates household savings decisions, and hence the tradeoff between time investments in children now and cash investments later in life. Abbott et al. (2016) focuses on the interaction between parental investments, state subsidies and education decisions, but abstract away from the role of parents in influencing ability prior to the age of 16. Castaneda et al. (2003) and De Nardi (2004) build overlapping-generations models of wealth inequality that includes both intergenerational correlation in human capital and bequests, but does not attempt to model the processes underpinning the correlation in earnings across generations.

The rest of this paper proceeds as follows. Section 2 describes the data, and documents descriptive statistics on ability, education and parental investments. Section 3 lays out the dynastic model used in the paper. Section 4 outlines our estimation approach while, Section 5 then provides some reduced-form evidence on the impact of parental investments, before Section 6 provides some initial results on the relative importance of different channels in explaining intergenerational correlations in education, earnings and welfare. Section 7 concludes, and draws out some implications for policy.

2 Data and descriptive statistics

The key data source for this paper is the National Child Development Study (NCDS). The NCDS follows the lives of all people born in England, Scotland and Wales in one particular week of March 1958. The initial survey at birth has been followed by subsequent follow-up surveys at the ages of 7, 11, 16, 23, 33, 42, 46, 50 and 55.[1] During childhood, the data includes information on a number of ability measures, measures of parental time investments (discussed in more detail below) and parental income. Later waves of the study record educational outcomes, receipt of inter-vivos transfers, demographic characteristics, earnings and hours of work. For the descriptive analysis in this section, we focus on those individuals for whom we observe both their father's educational attainment (age left school) and their own educational qualifications by the age of 33. This leaves us with a sample of 9,436 individuals.

The main limitation of the NCDS data currently available for our purposes is that we do not have data on the inheritances received or expected by members of the cohort of interest. We therefore supplement the NCDS data using the English Longitudinal Study of Ageing (ELSA). This is a biennial survey of a

[1] The age-46 survey is not used in any of the subsequent analysis as it was a telephone interview only, and the data are known to be of lower quality

representative sample of the 50-plus population in England, similar in form and purpose to the Health and Retirement Study (HRS) in the US. The 2012-13 wave of ELSA recorded lifetime histories of inheritance receipt, and since we also observe father's education in those data, we can use those recorded receipts to augment our description of the divergence in lifetime economic outcomes by parental background. We focus on individuals in ELSA born in the 1950s, leaving us with a sample of 3,001.[2]

In the rest of this section, we document the evolution of inequalities over the lifecycle, and in particular how they relate to parental background and parental investments over time.

2.1 Ability and time investments

Table 1 shows the multiple measures of both ability and investments available in our data at ages 0, 7, 11, and 16. These multiple measures are a key advantage of our data because each of the measures likely has measurement error, and we can use the measurement error models described below to address these concerns.

Table 1: List of all measures used

Ability measures	*Investment measures*
Age 0:	
gestation	parents' interest in education (**mother** and father)
birthweight	outings with child (mother and father)
index composed of:	read to child (mother and father)
walking, talking, late respiration	father's involvement in upbringing
	parental involvement in child's schooling
Age 7:	
reading score	parents' interest in education (**mother** and father)
maths score	outings with child (mother and father)
drawing score	father's involvement in upbringing
copying design score	parents' ambitions regarding child's educational attainment
	library membership of parents
Age 11:	
reading score	parents' interest in education (**mother** and father)
maths score	involvement of parents in child's schooling
copying design score	parents' ambitions regarding child's educational attainment
Age 16:	
reading score	
maths score	

All investment measures are retrospective, so age 0 investments are measured at age 7, age 7 investments are measured at age 11, age 11 investments are measured at age 16.
Normalizing measures in **bold**.

Although our measurement error models combine multiple measures of to create indexes of ability

[2]The next wave of the NCDS, which will be in the field next year, is currently planned to collect information on lifetime inheritance receipt. We hope to use these new data in later versions of this work

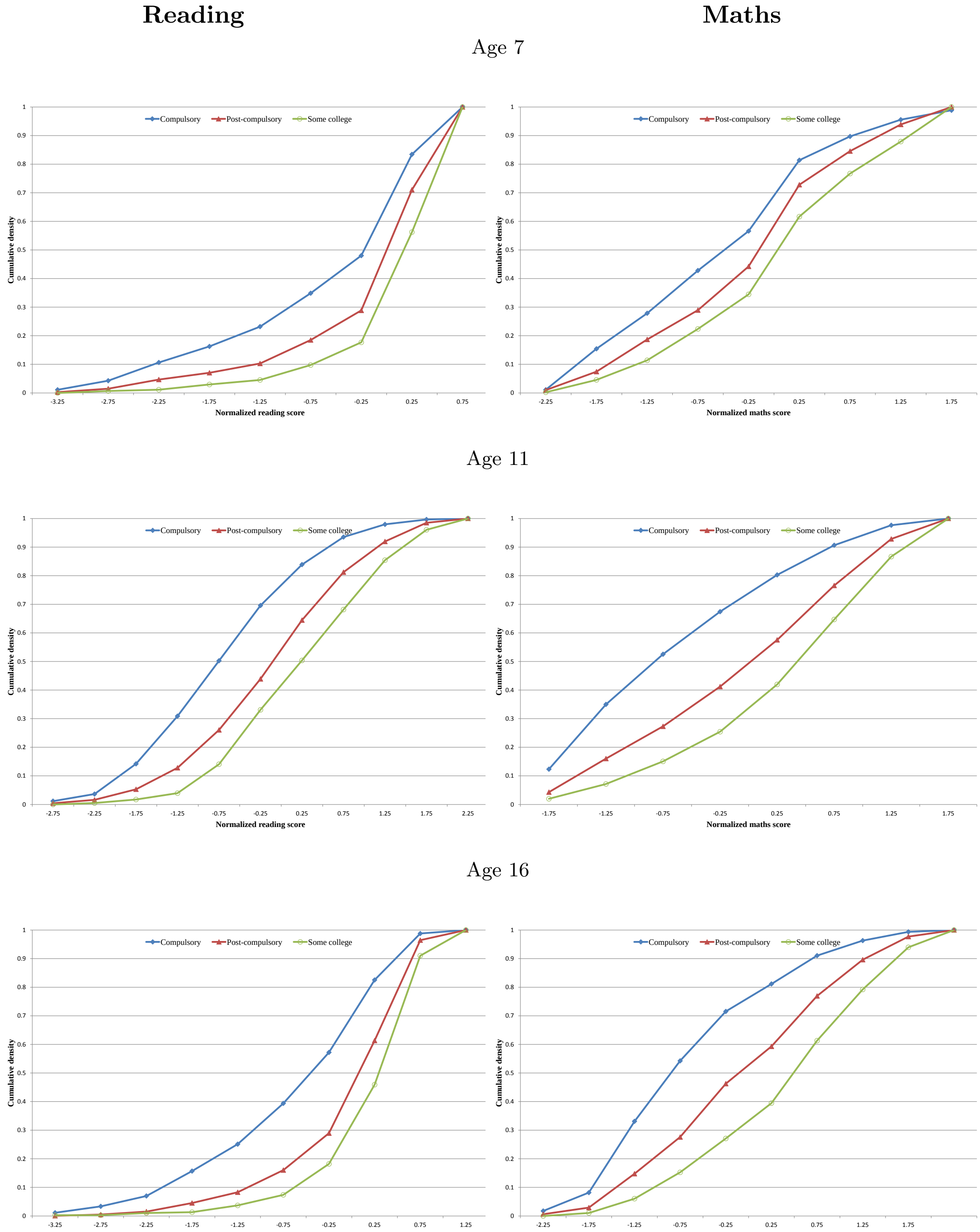
and investment, here we highlight some of the key features available in the raw data. Importantly, we show how these measures correlate with father's education to show evidence on the the intergenerational correlation of ability in the data and its drivers.

Figure 1 shows the cumulative distribution of normalised reading and maths ability at each age, splitting the sample according to father's education (compulsory only, some post-compulsory, some college - the proportion of children in each group is shown in the Appendix Table). For this age group of fathers, compulsory education roughly corresponds to leaving school at age 14, post-compulsory means leaving school between ages 15 and 18, and some college means staying at school until at least age 19. It shows that, as one might expect, children whose father has a higher level of education have higher ability; at the age of 7, 35% (35%) of the children of low-education fathers had a reading (math) score around one standard deviation or more below the mean, compared to just 8% (17%) of the children of high-education fathers. Similarly, 18% of the children of high-education fathers had a maths score around one standard deviation or more above the mean, compared to 8% of the children of low-education fathers.[3]

The second key thing to note from Figure 1 is that ability gaps by father's education widen through childhood. At the age of 7, 34% of the children of low-education fathers have above-average reading scores compared to 62% of the children of high-education parents - a gap of 28 percentage points. By age 11, that gap has widened to 34 percentage points, and by age 16 it stands at 39 percentage points. These gaps widen for maths also.

[3]For reading, a large number of students obtained a perfect score.

Figure 1: Reading and maths skills, by parental education



Tables 2, 3 and 4 provide some descriptive evidence that at least some of the widening in ability gaps by parental characteristics between ages 7 and 16 age can be explained by differential parental investments (we investigate this hypothesis more formally in Section **??**). Table 2 documents parental responses to a question about reading with their child, asked when the child is 7. It shows relatively small but potentially important differences in the frequency with which both mothers and fathers read to their children, splitting families according to the education of the father. For example, 34% of fathers with only compulsory education read to their 7-year-old children each week, compared to 53% of fathers with some college education.

Tables 3 and 4 present the child's teachers assessment of parental interest in the child's education, at the ages of 7 and 11 respectively. The differences by father's educational attainment are perhaps even more striking than those in reading patterns. When the child is 7, fathers with some college education are three times more likely to be judged by the teacher to be 'very interested' in their child's education as fathers with just compulsory education (65% compared to 22%). At the age of 11, the gap in paternal interest is very similar, with 72% of college-educated fathers judged to be 'very interested' in their child's education, compared to 25% of fathers with just compulsory education. The tables also show that having a higher-educated father dramatically reduces the risk of a child having parents with little interest in their education. Among those with a college-educated father, only around 10% have a mother or father who is judged to show 'little interest' in their education at the age of 11. On the other hand, among those whose father has only compulsory education that figure rises to around a quarter of mothers and nearly half of fathers.

Table 2: Frequency with which parents read to age-7 children

		Father reads...		
		Never	Sometimes	Every week
Father's education	Compulsory	30%	36%	34%
	Post-compulsory	20%	35%	45%
	Some college	18%	29%	53%
		Mother reads...		
		Never	Sometimes	Every week
Father's education	Compulsory	16%	37%	47%
	Post-compulsory	12%	31%	57%
	Some college	10%	23%	67%

Table 3: Teacher assessment of parental interest in education of age-7 child

		Father		
		Very interested	Some interest	Little interest
Father's education	Compulsory	22%	24%	55%
	Post-compulsory	44%	22%	34%
	Some college	65%	15%	20%
		Mother		
		Very interested	Some interest	Little interest
Father's education	Compulsory	35%	43%	23%
	Post-compulsory	60%	30%	10%
	Some college	76%	18%	6%

Table 4: Teacher assessment of parental interest in education of age-11 child

		Father		
		Very interested	Some interest	Little interest
Father's education	Compulsory	25%	29%	46%
	Post-compulsory	54%	25%	21%
	Some college	72%	16%	12%
		Mother		
		Very interested	Some interest	Little interest
Father's education	Compulsory	35%	38%	26%
	Post-compulsory	61%	27%	12%
	Some college	76%	16%	8%

2.2 Educational attainment and school type

Table 5 shows the correlation in educational attainment between fathers and their children. It shows two dramatic impacts of paternal education on educational outcomes. First, having a high-educated father makes it much less likely that a child will end up dropping out of high school.[4] 30% of the children of fathers with just compulsory education end up as high-school dropouts, compared to only 10% of those whose fathers have some post-compulsory education, and just 2% of those whose father have some college education. Second, having a high-educated father makes it much more likely that a child will end up with some college education. Fully 66% of the children of college-educated fathers also end up with some college education, compared to only 20% of those whose fathers only have compulsory education.

[4] In the UK context, we define 'high school dropout' as not having any of the academic qualifications obtained at age 16 (formerly O-Levels, now GCSEs)

Table 5: Intergenerational correlation in education

		Child's education		
		High-school dropout	High-school graduate	Some college
Father's education	Compulsory	30%	50%	20%
	Post-compulsory	10%	47%	43%
	Some college	2%	32%	66%

Of course, it is in theory possible that all of the intergenerational correlation in education is explained by the relationship between parental education and ability documented in the previous sub-section. However, one might also think that differences in the quality of the schools attended by children from different backgrounds also plays a role. In our particular institutional context (children attending high school in Britain in the late 1960s and early 1970s) a key dimension in which schools differed in quality was their 'type'. The majority of children attended 'comprehensive' or 'secondary modern' public schools which drew their students from across society (henceforth we refer to this type of school simply as comprehensive). A small proportion attended 'grammar' schools: public schools to which admittance was by an ability test at the age of 11. In addition to the peer effects associated with attendance at such a school, these grammar schools attracted much better teachers on average, and were much more focused on university (college) attendance than other public schools. Finally, a small minority of children went to private schools.

Table 6 shows the distribution of children across these three different types of school. As one might expect, those with higher educated fathers are dramatically more likely to have attended higher quality schools. 30% of those whose fathers went to college attended a private high school compared to just 2% of of those with low-educated fathers, and a further 26% attended a grammar school, compared to just 9% of those with low-educated fathers. Of course some of this discrepancy (particularly in the case of grammar schools) might well be accounted for by the differences in ability documented above, but they also reflect differential financial investments in children's education. The most obvious form of educational investment is paying for private education, which is much higher quality on average than public education. However, educational investments could also take less direct forms, such as paying the house price premium associated with living in the neighborhood of a good public school.

Table 6: High-school type by father's education

		Child's school type		
		Comprehensive	Grammar	Private
Father's education	Compulsory	89%	9%	2%
	Post-compulsory	68%	18%	13%
	Some college	43%	26%	30%

These financial investments differ from inter-vivos transfers and bequests in terms of timing, but also more importantly in that they directly impact on children's earnings, through the returns to education. The relationship between school type and educational attainment, conditional on ability, is shown by Figures 2 and 3. We divide our sample into quintiles of ability (measured at age 16 in the way described above), and then plot the probability of completing high school and attending college respectively separately for individuals that attended each of the three school types.

Figure 2 shows that at all levels of ability outside the top quintile, children who attend a grammar or private school are much more likely to complete high school than those of the same ability attending a normal public school. For example, 80% of children in the middle quintile of age-16 ability at a comprehensive school complete high school, compared to around 95% of those of the same ability who attend either a grammar or private school.

Figure 3 shows that attendance at private school provides a clear boost to the probability of college attendance conditional on age-16 ability. While individuals in the middle quintile of ability who attend a comprehensive school have less than a 30% chance of ending up with some college education, those with the same ability who attended private school have more than a 40% chance.

Figure 2: The impact of school type on completing high school

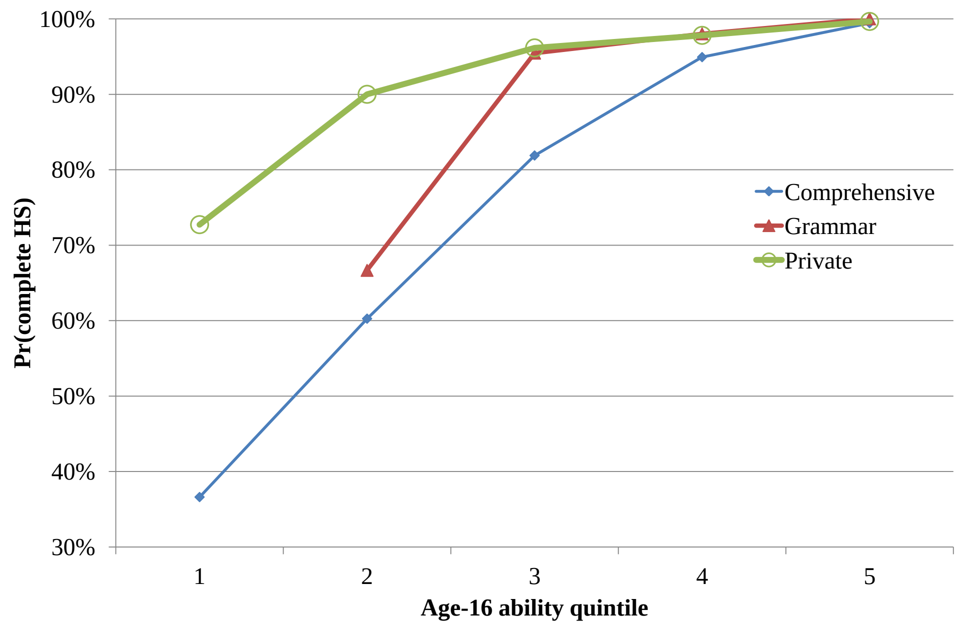


Figure 3: The impact of school type on attending college

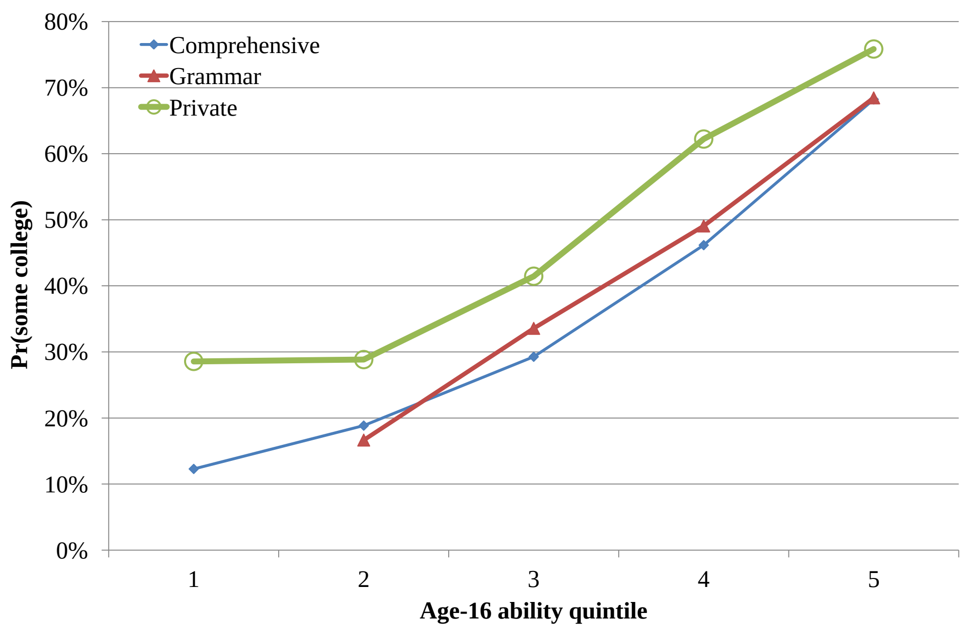


Table 7: Receipt of inter-vivos transfers and bequests by father's education

		Inter-vivos transfers by age 33		
		Mean (£)	Received	Mean exc. zeros (£)
Father's education	Compulsory	5,805	24%	24,281
	Post-compulsory	11,071	41%	27,008
	Some college	31,547	55%	56,933
		Inheritances (1950s birth-cohort)		
		Mean (£)	Received	Mean exc. zeros (£)
Father's education	Compulsory	17,180	26%	66,545
	Post-compulsory	43,901	40%	110,024
	Some college	55,669	46%	120,843

2.3 Inter-vivos transfers and bequests

Table 7 documents the receipt of inter-vivos transfers and bequests of the NCDS cohort so far, again splitting by father's education. As explained at the start of this section, the top panel draws on the NCDS data itself, while the bottom panel uses ELSA data instead, as information on inheritance receipt is not yet available in the NCDS.

The table shows that inter-vivos transfers are a significant source of economic resources for young adults, and that as one would expect are much more significant for those with higher-educated parents. By the age of 33, 55% of those whose fathers attended college had received an inter-vivos transfer, of an average of around £50,000. While this is the mean of a highly right-skewed distribution, these figures indicate an important role for inter-vivos transfers relieving borrowing constraints in this part of the lifecycle. At the same age, 24% of those with low-educated fathers had received an inter-vivos transfer, of an average size of just under £25,000.

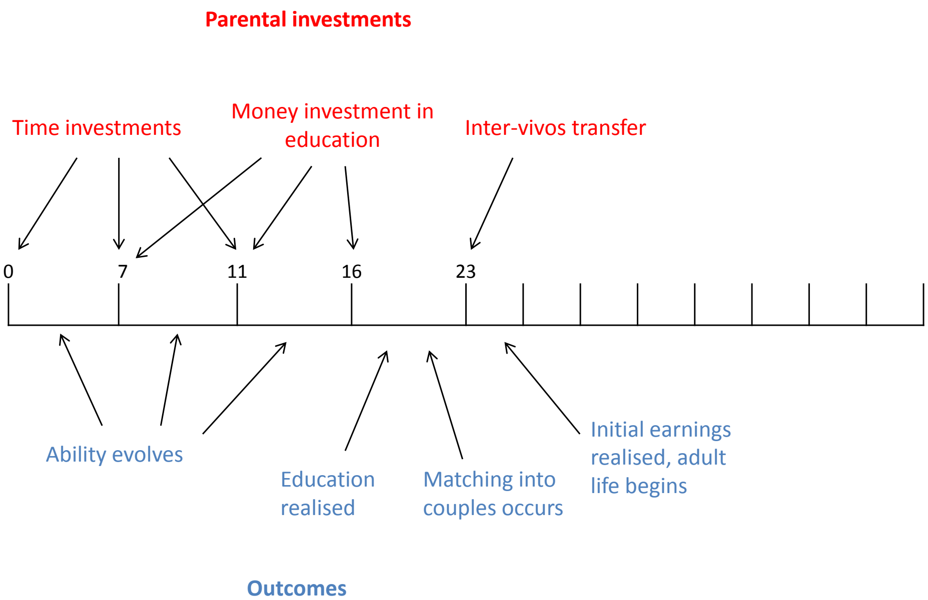
Evidence from ELSA data suggests that differences in inheritance receipt by parental background are also significant. 46% of those with college-educated fathers have received an inheritance, compared to 26% of those with low-educated fathers, and among those who have received an inheritance, those with college-educated fathers have received around twice as much on average (£120,843 compared to £66,545) . The net result is that those with college-educated fathers have inherited around £40,000 more than those with low-educated fathers. This is likely to understate the true difference in mean lifetime inheritance receipt between these groups; some of those born in the 1950s will still have living parents, and differential mortality means it is in fact likely that this applies to a larger share of those with high-educated fathers.

3 Model

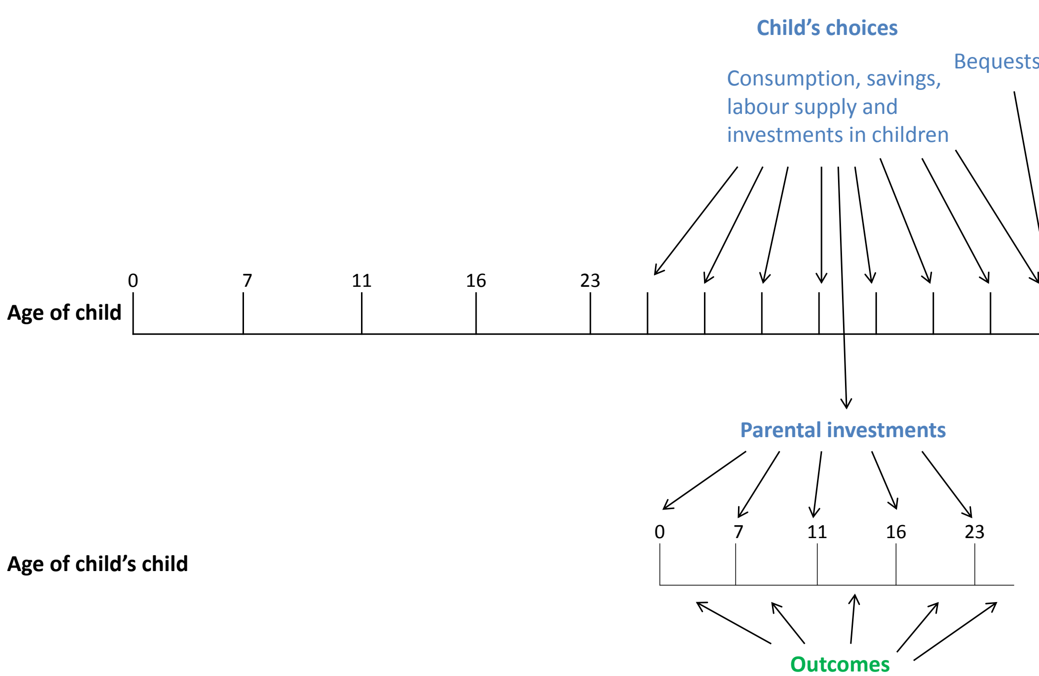
This section describes a dynastic model of consumption and labor supply in which parents can make different types of transfers to their children. The model can be used to a) evaluate how particular intergenerational transfers affect household behavior, b) compare the relative insurance value of these types of transfers and c) simulate household behavior and welfare under counterfactual policies (for example, under reforms to estate taxation). Figure 4 provides an overview of the dynastic model. During childhood, parental time investments in children and money investments in education affect the evolution of the child's ability and their educational attainment. Children are then matched in couples, receive any inter-vivos transfer from their parents and begin adult life. They then have their own children, and alongside the standard choices of consumption and labour supply they choose how much to invest in their own children, with implications for their children's future outcomes.

Figure 4: The Life Cycle of an Individual

(a) Childhood



(b) Adulthood and Investment in Children



We now provide formal details of the model. First, we outline a production function for ability, schooling and education in Section 3.1. We then outline the decision problem of a couple with a dependent child in Section 3.2.

3.1 A production function for ability, schooling and education

This section describes the production function for ability, schooling and education from ages birth to age 23. Over this part of the life cycle, the child makes no decisions. However, their parents do make decisions about the investments of time and goods received by their children. These choices do not directly impact the contemporaneous utility of the child, but leads to higher higher wages, incomes and higher quality spouses later in life, which does increase their later life outcomes.

3.1.1 Child ability production function

A child's ability of generation 1 at birth is given by:

$$ab_0^1 = f_{ab_0}(ab_0^0, u^{ab_0}) \tag{1}$$

where ab_0^1 is initial ability of the generation 1 child at birth, and ab_0^0 is the initial ability of the children's parents and $u_0^{ab_0}$ is a stochastic variable that generates heterogeneity in initial ability, conditional on parental education. This equation allows for an intergenerational correlation of initial ability that does not depend on choices, allowing us to create a measure of "nature".

Between birth and age 16, child ability updates each period according to the transition equation:

$$ab_{t+1} = f_{ab}(ab_t, ed^m, ed^f, ti_t^m + ti_t^f, st_t, u_{t+1}^{ab}) \tag{2}$$

The rate of growth of a child's ability depends on his/her parents' level of education, ed^m and ed^f (where m and f index male and female respectively), the sum of the time investments (ti_t^m, ti_t^f) those parents make, and the child's school type (st). There is also a stochastic component to the ability transition equation (u_{t+1}^{ab}). Ability evolves until the age of 16, after which it does not change (ab_{16} without a subscript denotes final ability).

3.1.2 School type production function

School type (st) is assumed not to vary between the start of education and the age of 11. At the age of 11, school type is realised as one of three outcomes: 1) Private $(st = p)$, 2) Public – high quality $(st = g)$,

3) Public – low quality ($st = m$).[5]

Parents can make one or both of two types of money investments in their children's schooling. First, they can pay a quantity of their choosing (mi^g) to attempt to get their children into a high quality public school (one can think of this is paying a premium to locate in a district where access to good quality schools is easier). Second they can make money investments in private schooling, paying a cost ($mi^p = p$), to guarantee that their child gets into a private school. We model the outcome of the school type as following a two-stage process. First, the child's type of public school is realised (st^g is a binary indicator of getting an *offer* at a high quality - or 'grammar' school). This is a stochastic function of the child's ability, parents' education and parent's choosing to spend money living (mi^g) in a location where access to good schools is easier:

$$st^g = f_{st^g}(ab_{11}, ed^m, ed^f, mi^g, u^{st^g}) \tag{3}$$

Second, after observing the type of public schooling on offer for their child, parents decide whether or not to pay for private schooling. They can accept the public option that their child has been given and pay $mi^p = 0$ or to reject it and pay $mi^p = p$ for private schooling.

This process can be summarised as follows:

$$st = \begin{cases} m \ \text{ if } \ mi^p = 0 \text{ and } st^g = 0 \\ g \ \text{ if } \ mi^p = 0 \text{ and } st^g = 1 \\ p \ \text{ if } \ mi^p = p \end{cases}$$

Total money investments $mi = mi^g + mi^p$: the sum of payments aim at gaining access to good public schools and those aimed at securing access to private schools.

3.1.3 Education Production Function

Education takes one of three values: High School drop out, High School graduate and Some College. It is realised in the period prior to a child turning 23. The education production function depends on the (now grown-up) child's ability, their school quality and a stochastic variable (u^{ed}).

$$ed = f_{ed}(ab, st, u^{ed})$$

[5]This component of the model is motivated by the institutional structure that faced the cohort represented by our main data. For this cohort, children took an exam at the age of 11 (the 'eleven-plus'). Children who performed well in this exam got a place in a selective 'grammar school'. Children who performed less well got a place in a 'secondary modern' or 'comprehensive' school. In our counterfactual analysis, we will explore scenarios in which there is no link between ability and the quality of public schooling.

3.2 Parents' decision problem

3.2.1 Stages of life

An individual's adult lifecycle starts at the age of 23 at which point their ability has been formed through their parents' decisions, their education has been realised and they have been matched into couples. Their lifecycle has three stages. First, there is the **early adult phase**, from the age of 23 to 48 when couples make decisions as a collective unit and have a dependent child. There is then a one-period **transition phase** at the age of 49 which is the last age at which they make decisions on behalf of their child. From the age of 50, their child has grown up and they enter their **late adult phase**. During this phase couples are subject to stochastic mortality risk.

In outlining the dynastic model we describe below a lifecycle decision problem of a single generation. All generations are, of course, linked; each member of the couple whose decision problem we specify has parents, and they, in turn, will have children. We will refer to the generation whose problem we outline as generation 1, their parents as generation 0, and their children as generation 2. In the exposition below, model periods are indexed by the age of the members of the couples in generation 1.[6]

3.2.2 Initial conditions and marital matching

Individuals start the decision-making phase of their life in couples at the age of 23. Individuals differ at the start of life in their ability, their level of education and their initial wealth. The first two are generated according to the production functions with inputs determined endogenously by their parents. The third – initial wealth – will come as a cash gift from their parents (the parents' decision problem is outlined below).

Before they make any decisions, individuals are matched into couples and acquire a dependent child at the age of 26. There is probabilistic matching between men and women which is based on education and ability. The probability that a man of education ed^m gets married to a women with education ed^f is given by $Q^m(ed^m, ed^f)$. The (symmetric) matching probabilities for females are $Q^f(ed^f, ed^m)$. Everyone is matched into couples – there are no singles in the model.

3.2.3 Utility and Demographics

The utility of each member of the couple $g \in \{m, f\}$ (male and female respectively) depends on their consumption

[6] That is, subscripts are an index of calendar time, not of age For example, $V^1_{50}()$ is the value function of generation 1 at the age of 50, but $V^2_{50}()$ is the value function of generation 2 in the year that generation 1 was 50 years old.

and leisure:

$$u_g(c,l) = \frac{(c^{\nu_g} l^{(1-\nu_g)})^{1-\gamma}}{1-\gamma}$$

We allow the relative preferences for consumption and leisure to vary with gender. Household preferences are given by the equally-weighted sum of male and female utility:

$$u(c, l^m, l^f) = u_m(c^m, l^m) + u_f(c^f, l^f)$$

and the consumption outcome is efficient within the household.

Mortality is stochastic - the probability of survival of a couple (we assume that both members of a couple die in the same year) to age $t+1$ conditional on survival to age t is given by s_{t+1}. We assume that death is not possible until the household enters the late adult phase of life at the age of 50 and that death occurs by the age of 110 at the latest.

3.2.4 Discounting and intergenerational altruism

In discounting their future utility each generation applies a discount factor (β). Each generation is altruistic regarding the utility of their offspring (and indeed future generations). In addition to the time discounting of their children's future utility (which they discount at the same geometric rate at which they discount their own future utility), they additionally discount it with an intergenerational altruism parameter (λ).

3.2.5 Decision Problem in Early Adult Phase of Life

Decisions During this phase couples in generation 1, matched into couples are making decisions on their own behalf and on behalf of their dependent child (generation 2). They make up to four choices each period. These are (with the time periods in which those decisions are taken given in parentheses):

1. Household consumption – c (each period)

2. Hours of work for each parent – h^m, h^f where m and f index hours of work by the male and female respectively (each period). We allow each parent to work full-time, part-time or not at all.

3. Time investments in children – ti (up to and including the age at which their child turns 11)

4. Private schooling choice (equivalently money investments in children's education) – mi (only at the age that their child turns 11)

Constraints Parents face two types of constraints. The first is an intertemporal budget constraint at the household level

$$a_{t+1} = (1+r)(a_t + y_t - c_t - mi_t) \tag{4}$$

where a is parental wealth, y is household income and the other variables have been defined above. Wealth must be non-negative in all periods. The second constraint is a per-parent ($g \in \{m, g\}$) intratemporal time budget constraint:

$$T = l_t^g + ti_t^g + h_t^g \tag{5}$$

where T is a time endowment, l^g is leisure time and the other variables have been defined above.

Earnings and Household income Household income is given by $y = \tau(e^m, e^f)$, where $\tau(.)$ is a function which returns net-of-tax income and e^m and e^f are male and female earnings respectively. Earnings are equal to hours multiplied by the wage rate, e.g.: $e^f = h_t^f w_t^f$. That wage rate evolves according to a process that has a deterministic component which varies with age and a stochastic (AR(1)) component.

$$\begin{aligned} \ln\, w_t &= \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 \ln ab_{16} + \delta_5 PT + v_t \\ v_t &= \rho v_{t-1} + \eta_t \\ \eta &\sim N(0, \sigma^2) \end{aligned}$$

where PT is a dummy for working part time. While the associated subscripts are suppressed here, each of $\{\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \rho, \sigma^2\}$ varies by gender (g) and education (Ed).

Uncertainty In this phase couples face uncertainty over the innovation to their wage equation, over the stochastic innovations to the child ability production function and the school type production function. The joint distribution of these stochastic variables ($\mathbf{q}_t^e \equiv \{\eta_t^m, \eta_t^f, u_t^{ab}, u_t^{st}\}$) is given by $F_t^e(\mathbf{q}_t^e)$.

State Variables The vector of state variables for generation 1 during the early adult phase of life contains (collected in the vector $\mathbf{X}^{1,e}$):

1. Age (t)

2. Assets (a^1)

3. Wage rates ($w^{m,1}, w^{f,1}$)

4. Eduction levels ($ed^{m,1}, ed^{f,1}$)

5. Own abilities ($ab^{m,1}, ab^{f,1}$)

6. Child's gender (g^2)

7. Child's ability (ab^2)

8. Child's school type (st^2)

where we make explicit the generation to which the state variable corresponds.

Value function The value function for generation 1 in the early adult phase of life ($V^{1,e}$) is given below in expression (6):

$$\begin{aligned} V_t^{1,e}(\mathbf{X}_t^{1,e}) \quad &= \quad \max_{c_t, h^m, h^f, ti, mi} \left(u(c_t, l_t^m, l_t^f) + \beta \int V_{t+1}^{1,e}(\mathbf{X}_{t+1}^{1,e}) dF_{t+1}^e(\mathbf{q}_{t+1}^e) \right) \\ &s.t. \quad \text{i) the intertemporal budget constraint in equation (4)} \\ &\qquad\ \ \text{ii) and the time budget constraints in equation (5)} \end{aligned} \tag{6}$$

3.2.6 Decision Problem in the Transition phase

The final period in which a couple is making decisions on behalf of their dependent child is when they are 46 (and their child is 23).

Decisions During this phase couples, couples make three sets of choices:

1. Household consumption – c (each period)

2. Hours of work for each parent – h^m, h^f where m and f index hours of work by the male and female respectively (each period)

3. A cash gift (x) to their children. This gift represents inter-vivos transfers and inheritances.

Constraints Parents once again face two types of constraints – an intratemporal time constraint and and an intertemporal budget constraint. The former is the same as that given in equation 5 in describing the early adult phase of the lifecycle (except that time investments in children will now always be zero). The intertemporal budget constraint in this phase takes account of the cash gifts and is given in equation 7.

$$a_{t+1} = (1+r)(a_t + y_t - c_t - x_t) \tag{7}$$

State variables The set of state variables ($\mathbf{X}^{1,tr}$) in this phase is that same as in the early phase of adulthood ($\mathbf{X}^{1,e}$).

Uncertainty Couples now face two distinct types of uncertainty. The first is uncertainty regarding their own circumstances next year – that is their next period wage draws ($\mathbf{q}^{tr}_{t+1} \equiv \{\eta^m_{t+1}, \eta^f_{t+1}\}$ with distribution given by $F^{tr}_{t+1}(\mathbf{q}^e_{t+1})$). The second is uncertainty over the characteristics of their child the following period. The dimensions of uncertainty here are the child's education, their initial wage draw and the attributes of their future spouse (his/her ability, education level, assets, and initial wage draw). The stochastic variables are collected in a vector $\mathbf{p}_{t+1}$, and their joint distribution is given by $H()$.

Value function The decision problem of generation 1 in the transition phase of life is:

$$\begin{aligned} V^{1,tr}_t(\mathbf{X}^{1,tr}_t) \quad &= \max_{c_t,h^m,h^f,x} \Bigg(u(c_t, l^m_t, l^f_t) \quad +\beta \int V^{1,l}_{t+1}(\mathbf{X}^{1,l}_{t+1}) dF_{t+1}(\mathbf{q}_{t+1}) \\ &\qquad\qquad +\beta\lambda \int V^{2,e}_{t+1}(\mathbf{X}^{2,e}_{t+1}) dH(\mathbf{p}_{t+1}) \Bigg) \\ &\qquad s.t. \qquad \text{the intertemporal budget constraint in equation (7)} \\ &\qquad\qquad\qquad \text{and the time budget constraint in equation (5)} \end{aligned} \tag{8}$$

Note that there are two continuation value functions here. The first is the future expected utility of that the decision-making couple will enjoy in the next period (when they will enter the late adult phase). The value function (given in equation (9)) must be integrated with respect to next period's wage draws, which are stochastic, and discounted by β, the time discount factor. The second continuation value function is the expected value of the couple to which the child of the generation 1 decision-maker will belong to. The (altruistic) parents take this into account in making their decisions. This continuation utility is discounted by both the time discount factor and the altruism parameter (λ). This value function is the early adult value function for generation 2 (the equivalent for generation 2 of the value function given in equation (6)).[7]

3.2.7 Decision Problem in the Late Adult phase

At this stage the children of generation 1 have entered their own early adult phase and the generation 1 couple enters a 'late adult phase',

Decisions During this phase households make labor supply and consumption/saving decisions only.

Uncertainty There is uncertainty over their next period wage draws ($\mathbf{q}_{t+1}^{l} \equiv \{\eta_{t+1}^{m}, \eta_{t+1}^{f}\}$ with distribution given by $F_{t+1}^{l}(\mathbf{q}_{t+1}^{l})$ and there is now stochastic mortality (where assume that both members of the couple die in the same period).

State Variables The vector state variables ($\mathbf{X}^{1,l}$) during the late adult phase of life are the same as those for the early adult phase except that the (now-grown-up) child's ability is no longer a state variable):

1. Age (t)
2. Assets (a)
3. Wage rates (w^m, w^f)
4. Eduction levels (ed^m, ed^f)
5. Own abilities (ab^m, ab^f)

Value function The decision problem in the 'late adult' phase of life can be expressed as:

$$\begin{aligned} V_t^{1,l}(\mathbf{X}_t^{1,l}) \quad &= \max_{c_t^1, h^{m,1}, h^{f,1}} \left(u(c_t, l_t^m, l_t^f) + \beta s_{t+1} \int V_{t+1}^{1,l}(\mathbf{X}_{t+1}^{1,l}) dF_l(\mathbf{q}^l) \right) \\ s.t. \quad & \text{the intertemporal budget constraint in equation (4)} \\ & \text{and the time budget constraint in equation (5)} \end{aligned} \tag{9}$$

where s_{t+1} is the probability of surviving to period $t+1$, conditional on having survived to period t.

[7] Recall that the timing convention that we index all value functions in this exposition by the age of generation 1. That is, $V_{t+1}^{2,e}$ is the value function for generation 2 when generation 1 is aged $t+1$.

4 Estimation

We adopt a two-step strategy to estimate the model. In the first step, we estimate or calibrate those parameters that, given our assumptions, can be cleanly identified outside our model. In particular, we estimate the human capital and educational production function, the wage process, the education transitions, marital sorting, and mortality rates from the data.

In the second step, we estimate the rest of the model's parameters (discount factor, consumption weight for both husband and wife, risk aversion, and altruism parameters)

$$\Delta = (\beta, \nu_f, \nu_m, \gamma, \lambda)$$

with the method of simulated moments (MSM), taking as given the parameters that were estimated in the first step. In particular, we find the parameter values that allow simulated life-cycle decision profiles to "best match" (as measured by a GMM criterion function) the profiles from the data.

Because our underlying motivations are to explain sources of income and parental investments in children, we match hours of work for both husbands and wives and also household time spent with children, by age and education. Because we wish to understand study money as well as time transfers to children, we also match education expenditures on children, as well as cash inter-vivos transfers to children when the children are older. Finally, to understand how households value their own utility in the present versus future, we match wealth data, which should be informative of the discount factor.

In particular, the moment conditions that comprise our estimator are given by

1. Employment rates, by age and education, for both men and women
2. Mean annual work hours of workers, by age and education, for both men and women
3. Mean annual household time spent with children, by age and education
4. Lifetime expenditure on child's education, by parents education level
5. Lifetime expenditure on inter-vivos transfers
6. Wealth at age 50 and 55

We observe hours and investment choices of these individuals, and thus match data for these individuals for the following years 1981, 1991, 2000, 2008, and 2013: when they were 23, 33, 42, 50 and 55.

The mechanics of our MSM approach are as follows. We compute life-cycle histories for a large number of artificial households. Each of these households is endowed with a value of the state vector of ability and wages of both self and spouse, and also wealth.

We discretize the asset and also the ability and wage grids for both spouses and, using value function iteration, we solve the model numerically. This yields a set of decision rules, which, in combination with the simulated endowments and shocks, allows us to simulate each individual's assets, work hours and home investment hours, and mortality. We use the profiles of hours and investments in children to construct moment conditions, and evaluate the match using our GMM criterion. We search over the parameter space for the values that minimize the criterion. Appendix **??** contains a detailed description of our moment conditions, the weighting matrix in our GMM criterion function, and the asymptotic distribution of our parameter estimates.

4.1 Estimating the Model of Parental Investments and Human Capital Accumulation

In our data we cannot directly observe children's skills (ab_t) or parents investment (I). However we have multiple noisy measures of each. In our analysis we explicitly account for this measurement error, using an approach where we generalize the methods in Agostinelli and Wiswall (2016a) to account for these multiple measures. We show how to use multiple measures (as in Cunha and Heckman (2008), Cunha et al. (2010)) but using a simpler system GMM approach rather than maximum likelihood and filtering methods.

Following AW, we use a restricted translog production function

$$\ln ab_{t+1} = \gamma_{1,t} \ln ab_t + \gamma_{2,t} \ln ti_t + \gamma_{3,t} \ln ti_t \cdot \ln ab_t + \gamma_{4,t} ed^m + \gamma_{5,t} ed^f + u_{ab,t} \tag{10}$$

We assume independence of measurement errors and use the noisy measures to instrument for one another. An important extension, relative to AW, is that we use many possible combinations of input measures to instrument for one another.

We use the same methodology to estimate the function

$$\ln ti_t = \alpha_{1,t} \ln ab_t + \alpha_{2,t} \ln ed_f + \alpha_{3,t} \ln ed_m + \alpha_{4,t} \ln y_t + u_{ti,t} \tag{11}$$

This equation is not derived from the structural model, but is infomative of how changes in ability, parental education, and income impact time investments, addressing measurement error. Thus we have our model structural parameters match the estimated α parameters in equation 13 using an indirect inference procedure. See the appendix for details on the estimation and identification of the parameters in this section.

4.2 Estimating the Wage Equation, Accounting for Measurement Error in Ability and Wages

We estimate the wage equation laid out in Section 3:

$$\ln w_t^* = \ln w_t + u_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 lnab_{16} + \delta_5 PT_t + v_t + u_t \text{ where} \tag{12}$$

$$v_t = \rho v_{t-1} + \eta_t,$$

$$u_t \text{ is IID measurement error in wages}$$

for each gender and education group. We estimate the wage equation parameters in two steps.

In the first step we estimate the δ parameters, accounting for measurement error in $lnab_{16}$ using the measurement error framework of AW In the second step we estimate the parameters of the wage shocks $\rho, Var(\eta)$, and $Var(v_{23})$ using a standard error components model, accounting for measurement errror in both $lnab_{16}$ and u_t. See appendix XX for details.

4.3 Selection in Wages–French Correction?

By assumption, $ab_{t,m}$ is uncorrelated with v_t in the population when estimating the wage equation, but we only observe wages of workers. We will account for this using a French correction.

4.4 Imputing Time Spent with Children

The NCDS data includes high quality information on measures of parental investments in children, but does not include the model consistent measure of actual hours spent with children. In order to construct a measure of actual hours spent with children, we first use our model to construct an index of time spent with children. To convert the index into hours, we use time data from the United Kingdowm Time Use survey. We calculate rank percentiles of our time index in the NCDS and and the rank percentiles of hours in the time use data. We then match the hours measure in. For example, a household is predicted to be at the 75th percentile of the time spent with children distribution in the NCDS data, then to that household we assign that household the hours with children observed at the 75th percentile of hours in the time use data.

5 First Step Estimation Results

In this section we describe results from our first-step estimation, that we use as inputs for our structural model, and the outputs that we require our model to match. These first step inputs describe the determinants of investments in children, how those investments affect childrens' ability, and how ability impacts success in the labor market. In particular, we present estimates of how children's ability, as well as parental resources, affect investments in children. We also present estimates of the effect of parental time investments on children's ability, and how that ability in turn affects subsequent education and adult earnings. This exploits a key advantage of our data - that we measure for the same individuals their parents' investments, their ability and the value of that ability in the labour market.

5.1 The Determinants of Parental Investments in Children

In Section 2, we documented that higher-educated parents spend more time time reading to their children and show more interest in their educational progress. Here we exploit this variation to create a measure of the time investments of parents in children using the methods described briefly in Section 4.1 and in more detail in the appendix. Throughout in estimation we use a GMM estimator with a diagonal weighting matrix.

For our model of investments between ages 0 and 7 (which are measured at age 7 but to remain consistent with the timing of the model we denote at time period 0) we estimate equation (13) which we rewrite here as

$$\ln ti_0 = \alpha_{1,0} \ln ab_0 + \alpha_{2,0} \ln ed_f + \alpha_{3,0} \ln ed_m + \alpha_{4,0} \ln y_0 + u_{ti,0} \tag{13}$$

Our time index is measured in logs, and thus we can interpret our coefficients as being measured as elasticities.

We find that a 1% increase in parents' income increases the index of time spent with the age 0 child by .11%, and a one year increase in mother's and father's education increases time spent with the child by .039% and .034%, respectively.

The effect of child's ability and parents' income on investments grows with age. We find that parents reinforce children's skills - a 1% increase in child ability at age 7 is associated with a 0.35% increase in age 7 investments. Parental education again has a positive effect on age 7 investments. The effect of income on investments becomes stronger such that a 1% increase in income leads to a 0.16% increase in investments.

These estimates grow further at age 11: a 1% increase in child ability at age 11 is associated with a 0.68% increase in age 11 investments. The effect of income on investments becomes stronger such that a 1% increase in income leads to a 0.26% increase in age 11 investments. Interestingly, father's education has a larger impact than mother's education on our age 11 time index.

Qualitatively these results are robust to also including a number of other covariates into the equation, such as parental age, number children and birth order.

Table 8: Determinants of time investments.

	IV weights			*Diagonal weights*		
	Coeff	90% CI		Coeff	90% CI	
Investment equation age 0						
log ability	**0.012**	-0.002	0.027	**0.002**	-0.009	0.012
log parents' income	**0.110**	0.083	0.146	**0.072**	0.052	0.097
mum education	**0.039**	0.029	0.050	**0.032**	0.024	0.041
dad education	**0.034**	0.026	0.043	**0.025**	0.017	0.032
N=						
Investment equation age 7						
log ability	**0.346**	0.311	0.397	**0.383**	0.351	0.425
log parents' income	**0.160**	0.113	0.213	**0.135**	0.093	0.183
mum education	**0.057**	0.040	0.073	**0.062**	0.045	0.077
dad education	**0.038**	0.025	0.050	**0.052**	0.038	0.065
N=						
Investment equation age 11						
log ability	**0.677**	0.645	0.722	**0.612**	0.576	0.661
log parents' income	**0.255**	0.193	0.326	**0.197**	0.139	0.257
mum education	**0.046**	0.023	0.067	**0.049**	0.028	0.068
dad education	**0.053**	0.034	0.070	**0.052**	0.037	0.067
N=						

Notes: education of mother and father measured in years.
Parent's income is log after tax annual labor income, measured at age 16.
For the investment equation at age 0, we use ability measured at age 0.
For the investment equation at age 7, we use ability measured at age 7.
For the investment equation at age 11, we use ability measured at age 11.
GMM estimates. Confidence intervals are bootstrapped using 100 replications.

5.2 Initial Conditions and the Intergenerational Correlation of Initial Ability

5.3 The Effect of Time Investments on Ability

In Section 2 we documented that children of high educated parents do better in cognitive tests, and that the ability gaps between children of high and low educated parents grow over time. Here we combine the multiple test scores to create a measure of skills, and estimate a human capital production function using the methods described briefly in Section 4.1 and in more detail in the appendix. Similar to our approach to estimating the determinants of time investment, we use a GMM estimator with a diagonal weighting matrix.

We estimate equation (10) for ability at ages 7, 11, and 16. Given the timing we estimate the following equation for 7 year olds:

$$\ln ab_{i,7} = \gamma_{1,0} \ln ab_{i,0} + \gamma_{2,0} \ln ti_{i,0} + \gamma_{3,7} \ln ti_{i,0} \cdot \ln ab_{i,0} + \gamma_{4,0} ed_i^m + \gamma_{5,0} ed_i^f + u_{ab,i,0}$$

where ab is latent ability, ti is our (latent) measure of parental time, and ed^m and ed^f is education of mother and father, respectively. Our ability index is measured in logs, and thus we can interpret our coefficients as being measured as elasticities.

Table 9: Determinants of log ability.

	IV weights			*Diagonal weights*		
	Coeff	90% CI		Coeff	90% CI	
Production function age 7						
log ability	**0.064**	0.052	0.104	**0.063**	0.048	0.107
log investment	**0.225**	0.199	0.260	**0.251**	0.216	0.292
interaction	**-0.021**	-0.057	-0.012	**-0.029**	-0.074	-0.017
mum education	**0.051**	0.032	0.068	**0.065**	0.042	0.084
dad education	**0.048**	0.035	0.063	**0.055**	0.041	0.071
N=						
Production function age 11						
log ability	**0.935**	0.901	1.042	**0.883**	0.836	0.971
log investment	**0.092**	0.062	0.111	**0.129**	0.100	0.151
interaction	**0.039**	0.012	0.081	**0.078**	0.053	0.128
mum education	**0.040**	0.017	0.059	**0.034**	0.015	0.049
dad education	**0.036**	0.016	0.055	**0.052**	0.035	0.065
N=						
Production function age 16						
log ability	**1.150**	1.128	1.235	**1.163**	1.133	1.225
log investment	**0.133**	0.073	0.178	**0.127**	0.085	0.164
interaction	**-0.005**	-0.043	0.040	**-0.016**	-0.051	0.025
mum education	**-0.013**	-0.034	0.006	**-0.008**	-0.022	0.006
dad education	**0.006**	-0.014	0.021	**0.000**	-0.013	0.011
N=						

Notes: education of mother and father measured in years.
Interaction is the interaction between log ability and log investment.
GMM estimates. Confidence intervals are bootstrapped using 100 replications.

We estimate the relationship between age 7 ability as a function of age 0 ability, age 0 time investments, the interaction of ability and time investments, and mother's and father's education. Estimates are presented in Table 9. It shows that time investments have a significant effect on changes in ability over time, even after conditioning on background characteristics and initial ability. A one percent increase in time investments at age 0 raises age-7 ability by 0.225 percent, a one percent increase in time investments at age 7 raises age-11 ability by 0.09 percent, and a one percent increase in time investments at age 11 raises age-16 ability by 0.13 percent.

Ability is very persistent, especially at older ages.

Interestingly, the interaction between ability and investments is negative for age 7, but positive for age 11. This implies that whilst at young ages, investments are more productive for low-skilled children, at older ages, productivity is higher for the higher-skilled ones. The positive and statistically significant coefficients on the age 11 interactions terms indicates that the ability production function does in fact exhibit dynamic complementarity at this stage of childhood (as found by Cunha et al. (2010)).

Qualitatively these results are robust to also including a number of other covariates into the equation, such as parental age, number children and birth order.

5.4 The effect of ability on wages

In the dynastic model with intergenerational altruism laid out in Section 3 parents do not receive any direct return from their children having higher ability at the age of 16. Instead, they include their children's expected lifetime utility in their own value function, with a weight determined by the intergenerational altruism parameter λ. Hence parental investments in children's ability (both through time and money investments in education) will be driven by the return to ability in the labour market. Here we focus on the return to ability in the labour market, as measured by its impact on wages. We estimate the wage equation laid out in Section 3:

$$\ln\ w_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 lnab_{16} + \delta_5 PT + v_t \tag{14}$$

for each gender and education group. Of course ability has an important indirect impact on wages through its relationship with education, but it also has a direct impact on wages conditional on education. This is shown by Table 10, which plots the estimates of δ_4 for each gender and education group. The interpretation of these coefficients is that they are estimates of the log-point increase in wages associated with a log-point increase in age-16 ability, conditional on education.

Table 10: Log-point change in earnings for a 1 log-point increase in ability

	Male	Female
High-school dropout	0.16	0.20
High-school graduate	0.31	0.29
Some college	0.55	0.38

The table shows that, as one would expect, age-16 ability has a significant positive impact on wages conditional on education for all groups. Perhaps more interesting, it finds evidence of complementarity between education and ability in the labour market, particularly for men. While male high-school dropouts see only a 0.15 log-point increase in hourly wages for every additional log-point of ability, men with some college education see an average increase of 0.55 log-points in hourly wages for every additional log-point of ability.

5.5 Marital Matching Probabilities

5.6 Other First Stage Estimates and Calibrations

5.7 The Moment Conditions We Match

6 Results

In this section we present some findings on the quantitative importance of different investments and stages of childhood in explaining the intergenerational transmission of economic advantage. First, we conduct a simple 'back of the envelope' exercise in decomposing the difference in lifetime income between individuals from different parental backgrounds into the proportions explained by different channels of investment. This is limited in a number of ways discussed below, but provides powerful suggestive evidence about the sources of intergenerational transmission of inequality. Second, we use a simplified version of the model described in Section 3 to quantify the differences in expected lifetime utility by parental education, and to decompose those differences into the proportions explained by different channels. This provides a more comprehensive measure of the relative importance of different channels, at the cost of relying on the structure of the model.

6.1 Decomposing the difference in lifetime income by parental education

In this analysis, we quantify the difference in expected lifetime income (as defined below) across children with fathers of the three different education levels defined and discussed in Section 2: compulsory, some post-compulsory and some college.

6.1.1 Methods

In this analysis, we focus on male members of the NCDS cohort, and define lifetime income as the sum of gross earnings during prime working age (between the ages of 25 and 55), plus any cash transfers and bequests from parents.

Differences in cash transfers and bequests from parents can be directly observed in the NCDS and ELSA data respectively, as reported in Table 7. To calculate differences in prime-age earnings we proceed in two steps. First, we estimate the earnings equation given in Sections 3 and **??**. Second, we calculate in the NCDS data the distribution across education and ability levels of individuals with each level of father's education. By combining these two things, we can calculate expected lifetime earnings for each paternal education group.

Having calculated expected earnings for each paternal education group given the actual distributions of ability and education within each group, we then do the same calculation for three counterfactual distributions of ability and education across each paternal education group:

1. We predict the distribution of age-16 ability and education for each paternal education group conditional on age-7 ability. Differences in expected earnings across groups in this scenario reveal how much of observed differences in earnings by paternal education can be explained by the differences in ability at age 7 shown in the first panel of Figure 1.

2. We predict the distribution of age-16 ability in the absence of differences in school type, and the predict education solely on the basis on that counterfactual age-16 ability distribution. The difference between expected earnings in this scenario and the previous one captures the effects of the faster growth in ability between 7 and 16 for children of higher-educated fathers, at least some of which is explained by the higher level of parental time investments in those children (as shown by the analysis in Section **??**).

3. We use the actual distribution of age-16 ability, but predict the education distribution for each group on the basis of age-16 ability and school type, ignoring other factors. The difference between expected earnings in this scenario and the previous scenario captures the effects of schooling differences on lifetime earnings. The difference between expected earnings in this scenario and true expected earnings captures the effect on lifetime earnings of other drivers of educational outcomes besides ability and school type.

6.1.2 Results

Overall differences in expected lifetime income (as defined above) for men with different levels of paternal education are shown in the first row of Table 11. Those with mid-educated fathers have expected incomes more than £150,000 higher than those with low-educated fathers, and the gap between those with low-educated and high-educated fathers is almost exactly £300,000. For reference, the lifetime income of those with low-educated fathers is a little over £850,000.

The rest of Table 11 decomposes these differences into distinct contributing factors.

- The first row of the decomposition shows differences in lifetime earnings in the first counterfactual scenario described above (age-16 ability and education predicted on the basis of age-7 ability). It shows that around 40% of the differences in lifetime income can be explained by differences in age-7 ability.

- The second row shows the difference between the first two counterfactual scenarios described above. It reveals faster growth in ability between 7 and 16 (not explained by different school types) explains around £23,000 of the gap between the children of low and mid-educated fathers, and around £39,000 of the gap between the children of low- and high-educated fathers (around 15% of the total gap in both cases).

- The third row shows the difference between the second and third counterfactual scenarios - schooling differences. The fact that those with higher-educated fathers are more likely to have attended private or grammar schools explains a little under 10% of the total differences in lifetime income

- The fourth row shows the difference between the final counterfactual scenario and actual expected earnings for each group. It suggests that differences in educational attainment conditional on ability and school type (explained by, for example, the role of financial support from parents) explains nearly 20% of the total gap in lifetime incomes across those from different parental backgrounds. It is perhaps surprising that differences in educational attainment conditional on school type and ability are twice as important in explaining differences in lifetime income as differences in school type.

- The final row of the Table simply documents differences in average inter-vivos transfers and bequests across paternal education groups. It shows that around 20% of the differences in lifetime income across these groups are attributable to differences in transfers and bequests, rather than differences in earnings.

To summarise, the decomposition analysis suggests that around 40% of the difference in lifetime income across paternal education groups is attributable to differences in ability at age 7, around 40% by subsequent divergence in ability and different educational outcomes, and around 20% by inter-vivos transfers and bequests received so far. Thus, while inter-vivos transfers are important, most of the lifetime differences in lifetime income between children of low versus high education fathers are realized by the age of 16.

6.2 Decomposing the difference in expected lifetime utility by parental education

There are a number of limitations with a comparison of expected lifetime incomes across individuals with different levels of parental education. Perhaps the most significant is that what individuals care about is the ex-ante difference in expected welfare, or expected utility. In this section we use a simplified version of the model laid out in Section 3 to estimate ex-ante expected lifetime utility for children with each level of parental education, expressed using compensating variation.

Table 11: Decomposition of differences in lifetime income by father's education

	Father's education	
	Some post-compulsory	Some college
Total difference	£156,000	£299,000
Accounted for by...		
Age-7 ability	£68,000	£115,000
Evolution of ability 7-16	£23,000	£39,000
School type differences	£11,000	£26,000
Attainment given ability and school type	£26,000	£58,000
Inter-vivos transfers and bequests	£28,000	£61,000

Memo: Lifetime income for those with low-educated fathers: £854,000

Notes: Differences relative to those with low-educated fathers (compulsory education only). Figures calculated for men.

6.2.1 Methods

The key simplification in the model used to estimate the results reported in this section is that we do not include intergenerational links - couples choose consumption and labour supply but not investments in their children. Hence the decision problem that households face corresponds to that described as the 'late adult phase' in Section 3. As a result, we do not explore how education and age-16 ability are determined within the model, but instead simply use the model to estimate expected lifetime utility given education and ability.[8] We calibrate the preference parameters of this simplified model to match labour supply and wealth moments.

Our approach then roughly follows that described in the previous section. We first use the model to estimate expected lifetime utility for each level of education and ability, and then combine that with the distribution of individuals from each parental background across education and ability to estimate the actual expected lifetime utility for each level of father's education. Then we can use the counterfactual distributions of education and ability discussed above to estimate expected lifetime utility for each parental education group in each of the counterfactual scenarios discussed. In order to provide a meaningful quantification of these differences in expected lifetime utility we calculate the consumption equivalent variation (CEV). This is the percentage increase in consumption in every state of the world required to make the children of less educated fathers indifferent between their ex-ante situation and that of those born to high-educated fathers.

[8]We also make a few further simplifications with respect to the model described in Section 3; namely, marital matching is on education only, we only allow individuals to choose whether to work full-time or not at all (no part-time choice), there are no earnings-related pensions (though each individuals receives a flat rate pension in retirement) and preference parameters are not gender specific.

6.2.2 Results

Table 12 shows the results from these CEV calculations. The first row of the table shows the total compensating variation required to make the children of low-educated fathers indifferent to being born to a mid-educated father (left-hand column) and a high-educated father (right-hand column). We estimate that the consumption of children of low-educated fathers in every state of the world would need to be increased by 6% for them to be indifferent with the children of mid-educated fathers, and by 12% for them to be indifferent with the children of high-educated fathers.

The rest of Table 12 decomposes these differences into distinct contributing factors.

- The first row of the decomposition shows differences in lifetime earnings in the first counterfactual scenario (age-16 ability and education predicted on the basis of age-7 ability). It shows that around 40% of the differences in expected lifetime utility (as measured by the CEV) can be explained by differences in age-7 ability. This is extremely similar to the proportion of the differences in lifetime income explained by age-7 ability in Table 11.

- The second row reveals faster growth in ability between 7 and 16 (not explained by different school types) explains between 15 and 20% of the differences in expected lifetime utility.

- The third row shows that schooling differences explain around 10% of the differences in expected lifetime utility - again a very similar result to that shown for expected lifetime income in Table 11.

- The fourth row shows that differences in educational attainment conditional on ability and school type (explained by, for example, the role of financial support from parents) explains around 30% of the total gap in expected lifetime utility across those from different parental backgrounds - a larger proportion than the 20% it explains of the difference in our measure of lifetime income. One potential reason for this is that the return to college attendance in the model is compounded by the existence of assortative matching in the marriage market - something not captured in our simple lifetime incomes analysis.

- The final row of the Table shows that inter-vivos transfers explain only 2% of the difference in expected lifetime utility between those with low- and mid-educated fathers, and only 7% of the difference between low- and high-educated fathers. This is somewhat in contrast to the findings presented in Table 11, which show that inter-vivos transfers and bequests explain around 20% of the differences across parental education in our measure of lifetime income. One reason for this difference is simply that bequests are not incorporated in the model, but perhaps more important is that these kinds of intergenerational transfers are extremely unequally distributed. Hence while they might have a meaningful impact on mean lifetime resources, from an ex-ante perspective they have very little effect on expected utility.

To summarise, this analysis of consumption equivalent variation (CEV) largely reinforces the conclusion of the previous analysis of expected lifetime income. Again, around 40% of the differences between those from different parental backgrounds can be explained by differences in ability by the age of 7, with the vast majority of the rest of the discrepancies being driven by later differences in ability and educational attainment, rather than cash transfers from parents to children. In fact, this analysis reinforces the conclusion that most of the meaningful differences between children of low versus high education fathers are realized by the age of 16.

Table 12: Decomposition of differences in expected lifetime utility by father's education

	Father's education	
	Some post-compulsory	Some college
Total compensating variation	5.9%	12.3%
Accounted for by...		
Age-7 ability	2.6%	4.8%
Evolution of ability 7-16	1.2%	2.0%
School type differences	0.6%	1.3%
Attainment given ability and school type	1.6%	3.9%
Inter-vivos transfers	0.1%	0.9%

7 Conclusions and policy implications

Understanding intergenerational links, and in particular the role of parental investments, is crucial for policymakers seeking to design redistributive tax and transfer policies that mitigate inequalities and improve social mobility, and wishing to understand the degree of intergenerational altruism (and hence the willingness of one generation to make sacrifices for another). In this paper we have documented substantial differences between children from different backgrounds in the evolution of cognitive ability through childhood, school quality and educational outcomes, and cash transfers received from their parents. A quantification of the implications of this differences for lifetime incomes suggests that around 40% of the gap between the sons of low- and high-educated fathers can be attributed to ability differences at the age of 7, a further 40% to subsequent differences in ability and educational attainment, and the final 20% to differences in the amount of inter-vivos transfers and bequests received. The relative importance of these different stages of life and forms of investment is also found when using a simple calibrated model to estimate the welfare gains from better parents (as measured using consumption equivalent variation). We provide evidence that at least some of the differences in ability and education are attributable to parental time investments in children and investments in school quality respectively.

All this has a number of implications for economic policy. At the most general level, the paper shows that policymakers interested in tackling the intergenerational transmission of inequalities need to consider policies designed to counter the inequality-increasing effects of each of the three forms of parental investment, since each proves to be quantitatively important in driving inequalities in income. Moreover, policymakers should bear in mind the substitutability of these different forms of investment - any attempt to shut down one channel of parental investments is likely to provoke a shift towards investment in other forms. In fact, the elasticity of substitution between these different forms of investment is a key determinant of the optimal policy response, and is something we will quantify in later versions of this paper.

A Appendix Table

Table 13: Proportion of children in each father's education group

	Compulsory	75%
Father's education	Post-compulsory	20%
	Some college	5%

B Our Adaptation of the Agostinelli-Wiswall Approach

B.1 Production Function

The production function for skills takes a restricted translog specification.

$$\ln ab_{t+1} = \gamma_{1,t} \ln ab_t + \gamma_{2,t} \ln ti_t + \gamma_{3,t} \ln ti_t \cdot \ln ab_t + \gamma_{4,t} ed^m + \gamma_{5,t} ed^f + u_{ab,t}$$

B.2 Parental Investment Function

Parental investments depend on a child's current skills, the parents' education and parental income. We use parameters of this reduced form policy function as target moments for the structural model.

$$\ln ti_t = \alpha_{1,t} \ln ab_t + \alpha_{2,t} \ln ed_f + \alpha_{3,t} \ln ed_m + \alpha_{4,t} \ln y_t + u_{ti,t} \tag{15}$$

B.3 Measurement

We don't observe children's skills (ab_t), or investments (ti_t) directly. However we observe m error-ridden measurements of each. These measurements have arbitrary scale and location. That is for each $\omega \in \{ab, ti\}$ we observe:

$$Z_{\omega,t,m} = \mu_{\omega,t,m} + \lambda_{\omega,t,m} \ln \omega_t + \epsilon_{\omega,t,m} \tag{16}$$

All other variables are assumed to be measured without error.

B.4 Assumptions on Measurement Errors and Shocks

Measurement errors assumed to be independent across measures and across time. Measurement errors are also assumed to be independent of the latent variables, household income and the structural shocks ($u_{ti,t}, u_{ab,t}, \ln ed_f, \ln ed_m, \ln y_t$).

B.5 Normalizations

As mentioned above, skills and investments do not have a fixed location or scale which is why we need to normalize them. For each period, we normalize the mean of the log-latent factor to be zero. Moreover, we set the scale $\lambda_{\omega,t,1} = 1$ for one normalizing measure $Z_{\omega,t,1}$.

In certain cases, AW have shown that these assumptions can lead to biases in the estimation of the coefficients. In our case, however, these normalizations will not lead to biased estimates as we use a production function which is consistent with a mean-zero process. Moreover, we do not have overidentification in the production function as we do not assume constant returns to scale. For more details, see Agostinelli and Wiswall (2016b).

B.6 Intial Conditions Assumptions

Period zero for us is at birth.

$$\mathbf{\Omega} = (\ln ab_0, \ln ti, ed_f, ed_m, \ln y_t) \sim N(\mu_\Omega, \mathbf{\Sigma}_\Omega) \tag{17}$$

$\mu_\Sigma = [0, 0, 0, 0, 0]$. The mean of $\ln ab_0$, $\ln ed_f$, $\ln ed_m$ and $\ln ti$ are 0 by normalization and without loss of generality.

B.7 Estimation

1. **Variance of latent factors** (the elements of $\mathbf{\Sigma}_\Omega$).

   Using equation (16) we can derive the variance of each of the latent factors:

   $$Cov(Z_{\omega,t,m}, Z_{\omega,t,m'}) = \lambda_{\omega,t,m}\lambda_{\omega,t,m'}\, Var(\ln \omega_t) \tag{18}$$

   Note that this is overidentified as there are many different combinations of m and m' that can be used here. Whilst AW arbitrarily choose one of the combinations, we use the bootstrap to estimate the variances of the objects in equation (18) and run a diagonal GMM in order to construct a unique $Var(\ln \omega_0)$. Because y, ed^m, ed^f are observable, it is straightforward to estimate the covariance of these with each other, as well as their covariance with ability and time investments.

2. **Scale parameters (λs) in skills measurement equations**. Here we estimate the parameters for the measurement equations for the child skill and investment latent variables. We have normalised $\lambda_{ab,t,1} = \lambda_{ti,t,1} = 1$ to set the scale of ab_t and of ti_t. For each other measure $m \neq 1$, and for $\omega \in \{ab_t, ti\}$, using equation (18) we can show that:

$$\lambda_{\theta_\omega,t,m} = \frac{Cov(Z_{\omega,t,m}, Z_{\omega,t,m'})}{Cov(Z_{\omega,t,1}, Z_{\omega,t,m'})} \tag{19}$$

Note that this is overidentified as there are many different combinations of m and m' that can be used here.

3. **Location parameters (μs) in skills measurement equations** We have normalized the mean of $\ln \theta_t$ and $\ln ti$ to zero. Therefore:

$$\mu_{\omega,t,m} = \mathbb{E}[Z_{\omega,t,m}] \tag{20}$$

4. **Calculation for next step** For each measure we need to calculate a residualized measure of each Z for $\omega_t \in \{ab_t, ti\}$:

$$\tilde{Z}_{\omega,t,m} = \frac{Z_{\omega,t,m} - \mu_{\omega,t,m}}{\lambda_{\omega,t,m}} \tag{21}$$

This will be used below in Step 1. Note that:

$$\ln \omega_t = \tilde{Z}_{\omega,t,m} - \underbrace{\frac{\epsilon_{\omega,t,m}}{\lambda_{\omega,t,m}}}_{\equiv \tilde{\epsilon}_{\omega,t,m}} \tag{22}$$

It gives log skills (or time input) plus an error rescaled to match scale of the skills (which is also the scale of skill measure 1).

5. **Estimate Latent Skill Production Technology**

We will only describe the estimation of the production technology, as the estimation of the investment equation is analogous. Recall the production function:

$$\ln ab_1 = \gamma_{1,0} \ln ab_0 + \gamma_{2,0} \ln ti_0 + \gamma_{3,0} \ln ti_0 \cdot \ln ab_0 + \gamma_{4,0} ed^m + \gamma_{5,0} ed^f + u_{ab,0}$$

and using equation (22) note that we can rewrite the above equation as:

$$\begin{aligned}\frac{Z_{ab,1,m}-\mu_{ab,1,m}-\epsilon_{ab,1,m}}{\lambda_{ab,1,m}} &= \gamma_{1,0}(\tilde{Z}_{ab,0,m}-\tilde{\epsilon}_{ab,0,m})+ \\ &\quad \gamma_{2,0}(\tilde{Z}_{ti,0,m}-\tilde{\epsilon}_{ti,0,m})+ \\ &\quad \gamma_{3,0}(\tilde{Z}_{ti,0,m}-\tilde{\epsilon}_{ti,0,m})\cdot(\tilde{Z}_{ab,0,m}-\tilde{\epsilon}_{ab,0,m})+ \\ &\quad \gamma_{4,0}ed^{m}+\gamma_{5,0}ed^{f}+ \\ &\quad u_{ab,0}\end{aligned} \tag{23}$$

or

$$\begin{aligned}\frac{Z_{ab,1,m}-\mu_{ab,1,m}}{\lambda_{ab,1,m}} &= \gamma_{1,0}\tilde{Z}_{ab,0,m}+ \\ &\quad \gamma_{2,0}\tilde{Z}_{ti,0,m}+ \\ &\quad \gamma_{3,0}\tilde{Z}_{ti,0,m}\cdot\tilde{Z}_{ab,0,m}+ \\ &\quad \gamma_{4,0}ed^{m}+\gamma_{5,0}ed^{f}+ \\ &\quad (u_{ab,0}-\tilde{\epsilon}_{ab,0,m}-\tilde{\epsilon}_{ti,0,m}-\tilde{\epsilon}_{ti,0,m}\cdot\tilde{\epsilon}_{ab,0,m}+\frac{\epsilon_{ab,1,m}}{\lambda_{ab,1,m}}).\end{aligned} \tag{24}$$

OLS is inconsistent here, as $\tilde{Z}_{ab,0,m}$ and $\tilde{\epsilon}_{ab,0,m}$ are correlated. We resolve this issue by instrumenting for $\tilde{Z}_{ab,0,m}$ using the other measures of ability $\tilde{Z}_{ab,0,m'}$ in that period.

Note that again, we have many possible input and output measures here, which help us identify the underlying coefficients. Rather than just picking one input and output measure as AW do, we use the whole set of relevant measures and estimate a system GMM with diagonal weights. By using the system GMM we make efficient use of all available measures.

6. **Estimate the Variance of the Production Function Shocks**

The variance of the structural skills shock can be obtained using residuals from equation (24), where $\pi_{\theta,0,m} \equiv (u_{ab,0}-\tilde{\epsilon}_{ab,0,m}-\tilde{\epsilon}_{ti,0,m}-\tilde{\epsilon}_{ti,0,m}\cdot\tilde{\epsilon}_{ab,0,m}+\frac{\epsilon_{ab,1,m}}{\lambda_{ab,1,m}}):$

$$Cov\left(\frac{\pi_{\theta,0,m}}{\lambda_{\theta,0,m}},\tilde{Z}_{\theta,0,m'}\right)=\sigma^2_{\theta,0,m}$$

As again, these covariances are overidentified, we use a bootstrap and diagonal GMM to estimate the shock variances efficiently. Again, the variance of the time investment shocks is estimated similarly.

C Estimating the Wage Equation, Accounting for Measurement Error in Ability and Wages

We estimate the wage equation laid out in Section 3:

$$\ln w_t^* = \ln w_t + u_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 lnab_{16} + \delta_5 PT_t + v_t + u_t \text{ where} \tag{25}$$

$$v_t = \rho v_{t-1} + \eta_t,$$

$$u_t \text{ is IID measurement error in wages}$$

and PT_t relates to part time status, for each gender and education group. We estimate the wage equation parameters in two steps.

In the first step we estimate the δ parameters by redefining:

$$\ln w_t^* = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 \tilde{Z}_{ab,16,m} + \delta_5 PT_t + v_t + u_t - \tilde{\epsilon}_{ab,16,m}$$

where $\tilde{Z}_{ab,16,m}$ is a normalized noisy measure of ability at 16 and $\tilde{\epsilon}_{ab,16,m}$ is measurement error in ability. Because $\tilde{\epsilon}_{ab,16,m}$ is correlated with measured ability $\tilde{Z}_{ab,16,m}$ by construction, we use the other measures of ability as instruments for $\tilde{Z}_{ab,16,m}$ to identify the δ coefficients.

In the second step we estimate the parameters of the wage shocks $\rho, Var(\eta)$, and $Var(v_{23})$

$$\widetilde{\ln w_t} \equiv \ln w_t^* - (\delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 \tilde{Z}_{ab,16,m} + \delta_5 PT_t) = v_t + u_t - \tilde{\epsilon}_{ab,16,m}$$

Note that we can estimate $V(\tilde{\epsilon}_{ab,16,m})$ from the production function

$$C(\widetilde{\ln w_t}, \widetilde{\ln w_{t+k}}) = \rho^k V(v_t) + V(\tilde{\epsilon}_{ab,16,m})$$

$$V(\widetilde{\ln w_t}) = V(v_t) + V(u_t) + V(\tilde{\epsilon}_{ab,16,m})$$

$$V(\widetilde{\ln w_{t+k}}) = \rho^k V(v_t) + \sum_{j=1}^{k} \rho^j V(\eta_{t+k}) + V(u_{t+k}) + V(\tilde{\epsilon}_{ab,16,m})$$

Note that ability at age 16 is constant so the variance also remains the same.

If we are willing to assume that $\rho = 1$, we could do a little better, since we could allow v_t to be correlated with other observables, and estimate the model using fixed effects.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. Violante (2016). Education policy and intergenerational transfers in equilibrium. Technical report, Cowles Foundation Discussion paper.

Agostinelli, F. and M. Wiswall (2016a, July). Estimating the technology of children's skill formation. Working Paper 22442, National Bureau of Economic Research.

Agostinelli, F. and M. Wiswall (2016b, 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.

Castaneda, A., J. Diaz-Gimenez, and J.-V. Rios-Rull (2003). Accounting for the u.s. earnings and wealth inequality. *Journal of Political Economy 111*(4), 818–857.

Caucutt, E. M. and L. Lochner (2012, October). Early and late human capital investments, borrowing constraints, and the family. Working Paper 18493, National Bureau of Economic Research.

Charles, K. K. and E. Hurst (2003). The correlation of wealth across generations. *Journal of Political Economy 111*(6), 1155–1182.

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.

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.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies 71*(3), pp. 743–768.

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

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

Fuster, L., A. Imrohoroglu, and S. Imrohoroglu (2007). Elimination of Social Security in a Dynastic Framework. *Review of Economic Studies 74*(1), 113–145.

Gayle, G.-L., L. Golan, and M. Soytas (2015). What is the source of the intergenerational correlation in earnings. Technical report, Mimeo.

Lee, S. Y. and A. Seshadri (2016). On the intergenerational transmission of economic status. *Unpublished manuscript, University of Wisconsin–Madison, Department of Economics*.

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.