Intergenerational Altruism and Transfers of Time and Money: A Life Cycle Perspective*

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

June 26, 2025

Abstract

Parental transfers of time, education, and money shape intergenerational mobility. Using panel data from birth to retirement, we estimate child skill production functions and embed them in a dynastic model of parental investment. We find that parental time investments have a dual purpose: investments increase children's human capital, but parents also enjoy time spent with children more than work. We then assess the effects of relaxing intergenerational borrowing constraints via student loans: student loans reduce persistence in earnings and education and raise parental welfare. However, while children newly able to attend college experience welfare gains, on average the higher debt levels reduce children's welfare.

*We are very grateful to Peter Arcidiacono, Richard Blundell, Judith Delaney, Andrew Hood, George-Levi Gayle, Limor Golan, James Heckman, Lance Lochner, Costas Meghir, Ananth Seshadri, and multiple seminar participants for helpful discussions. We are very grateful to Sree Ayyar, Jack Light and Alan Liu for excellent research assistance. Funding from the Social Security Administration through the Michigan Retirement Research Center (MRRC grant UM17-15) and the Economic and Social Research Council (Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (RES-544-28-50001), and NORFACE for Grant 462-16-120: Trends in Inequality: Sources and Policy, Grant: Inequality and the Insurance Value of Transfers Across the Life Cycle (ES/P001831/1), and Grant: Intergenerational Transfers, Insurance, and the Transmission of Inequality (ES/V001248/1)) for this work is gratefully acknowledged. O'Dea additionally acknowledges funding from the ESRC Secondary Data Analysis grant ref: ES/N011872/1. We are grateful to the Centre for Longitudinal Studies (CLS) for the use of these data and to the UK Data Service for making them available. The opinions and conclusions expressed are solely those of the authors and do not represent the opinions or policy of SSA or any agency of the Federal Government. Any errors are the authors' own.

†University of Bristol and Institute for Fiscal Studies

‡University of Cambridge and Institute for Fiscal Studies

§HEC Paris and Institute for Fiscal Studies

¶Yale University, NBER and Institute for Fiscal Studies

1 Introduction

The intergenerational persistence of education, earnings, and wealth is well documented[1]. Yet the role of parental preferences—over how they allocate time, financial resources, and educational support to their children—in shaping this persistence remains less well understood. This paper estimates a dynastic model with three key channels through which parents influence their children's outcomes: i) time investments in child and adolescent human capital formation, ii) financial support for education, and iii) transfers of wealth through inter-vivos gifts and bequests.[2] The model allows us to recover estimates of parental preferences and to assess how policy interventions, such as student loan programs, affect the welfare of both parents and their children.

We use data from the National Child Development Survey (NCDS), which is an ongoing panel of the entire population of Britain born in a particular week in 1958. The data set contains multiple measures of parental time investments and cognitive skill in childhood, as well as educational attainment and earnings during adulthood. We use these data to estimate child skill production functions in which parental time investments affect cognitive skill, which, together with education, determines wages. We then embed these production functions into a dynastic model in which altruistic parents choose consumption, labor supply, as well as transfers to their children of time, education, and cash.

We use the model for two main purposes. First, we estimate how parents value the time they invest in their children. Specifically, we ask whether this time is as costly as paid work, as rewarding as leisure, or somewhere in between. This allows us to estimate a novel component of preferences over an *input* to child skill formation, alongside the degree of parental altruism toward their children's *outcomes*. We find that time investing in children is neither completely viewed as leisure nor as work. Part of what drives this finding is that the returns to *parental* time investing in children are below the market wage. This finding contrasts with other evidence that the returns to *professional* time investing in children are above the market return (e.g., Heckman et al., 2010).

Second, we use these estimates to evaluate how relaxing intergenerational borrowing constraints—through access to student loans—affects the welfare of both parents and children. While such a policy improves the altruistic parents' welfare (since parents may choose not to use the loans), its effect on children is more nuanced. Looser borrowing constraints can enable greater educational investment, but may also result in children taking on debt to fund education that parents would otherwise have paid

[1]For evidence on intergenerational correlations, see Blanden et al. (2022) for education; Solon (1992), Dearden et al. (1997), Mazumder (2005), Chetty et al. (2014) for earnings; and Charles and Hurst (2003) for wealth.

[2]For evidence on parental time investments during childhood and adolescence and their impact on child development, see reviews by Cunha et al. (2006) and Heckman and Mosso (2014); for parental aid for education, see Belley and Lochner (2007), Abbott et al. (2019); and for cash gifts in the form of inter-vivos transfers and bequests, see Castaneda et al. (2003), De Nardi (2004).

for. We quantify the resulting impact on welfare for both parents' and children's generations. We find that student loans improve welfare of children who would not have attended college otherwise, while infra-marginal students, who would have gone to college regardless, bear new debt burdens without new benefits. These findings underscore a core tension in the design of programs which relax intergenerational borrowing constraints. Universally available loans promote mobility but risk imposing costs on the very cohorts they aim to help. Targeted, merit-based loans can improve average welfare, but do so by amplifying the advantages of some already well-off families. The generational welfare differences that we find may shine some light on the generational divide in attitudes to student debt (e.g., CNN, 2022).

Our model includes five mechanisms that contribute to intergenerational persistence in outcomes: i) borrowing constraints that limit college access for low-income families, ii) a positive correlation between parents' productivity in child skill production and their wages, iii) dynamic complementarities between time investments at different ages in the production of skills, and between skills and educational attainment in the production of wages, iv) positive assortative matching, which further reinforces persistence in household earnings, and v) cash transfers from parents to children. To our knowledge, this is the first paper to jointly model all these mechanisms. The estimated model produces an intergenerational earnings elasticity of 0.5, consistent with estimates in Dearden et al. (1997) and Blanden et al. (2007).

We have three main findings. First, in our estimated skill and earnings production functions, we find dynamic complementarities between early and later time investments in children. The complementarities between cognitive skill at age 16 and educational attainment in the wage process are particularly strong. Among men with a college education, a one standard deviation increase in log cognitive skill at age 16 is associated with 21 percent higher wages. For those with only compulsory schooling, the premium is just 8 percent.

Second, we estimate parental preferences over time investments. We find that, in order to rationalize observed behavior given the productivity of these investments, parental time investments must have a dual purpose. While investments increase children's human capital, parents also enjoy time spent with children more than work.

Finally, we use the estimated model to examine the effects of relaxing intergenerational borrowing constraints through student loans. Such a reform lowers the intergenerational persistence of education and earnings and increases the welfare of parents. For children, however, the average effect on welfare is negative. This masks substantial heterogeneity: children who would not have attended college under tighter borrowing constraints benefit from the reform, while those who would have attended regardless are worse off, as they incur higher levels of debt.

This paper relates to a number of different strands of the literature. The first is that which studies parental altruism and its role in driving intergenerational persistence in outcomes. The most closely related papers in that literature are those focused on the costs of, and returns to, parental investments in children. The four papers closest to ours in this literature are Caucutt and Lochner (2020), Lee and Seshadri (2019), Daruich (2022) and Yum (2022). Each of those papers, like ours, contains a dynastic model in which parents can give time, education, and money to their children. All four papers find that early life investments are key for understanding the intergenerational correlation of income. We build on the contributions of these papers in three ways.

The first is that we use the same sample throughout our analysis, enabling us to measure parental transfers, cognitive skill, and later life wage and other outcomes for one cohort of people in a single setting. This allows us to use the same cognitive skill measures for both estimation of the human capital production function and the wage equation, for example. We estimate the human capital production function and show directly how early life investments impact human capital and later life earnings. In contrast, previous papers lacked data that links investments at young ages to earnings at older ages, meaning that they have to calibrate key parts of the model. Our data linking early life investments to later life outcomes is key to estimating parental preferences over time investments, as it allows us to observe that the returns in terms of child human capital are not high enough by themselves to justify observed time with children.

The second is that we explicitly model the behavior of both men and women before and after they are matched into couples. This allows us to show the quantitatively important role that assortative matching plays in amplifying the impact of parental transfers in generating persistence in outcomes at the household level.

Finally, the focus of our paper is different. Caucutt and Lochner (2020) focus on identifying the role of market imperfections in rationalizing observed levels of parental investments. The aim of Lee and Seshadri (2019) is to simultaneously rationalize intergenerational persistence in outcomes and cross-sectional inequality in outcomes. Daruich (2022) focuses on the macroeconomic effects of large-scale policy interventions. Yum (2022) focuses on the role of heterogeneity in time investments. Our primary focus, facilitated by our data on the three parental inputs, is to understand parental preferences over transfers of time and money, and how those transfers interact with government policies and borrowing constraints in shaping intergenerational persistence in outcomes.

Other closely related papers include Del Boca et al. (2014), Gayle et al. (2022), and Mullins (2022), all of which develop models in which parents choose how much time to allocate to work in the labor market, leisure, and investment in children. These papers, however, do not include household savings decisions, and hence the trade-off between time investments in children now and cash investments later in life.

Castaneda et al. (2003) and De Nardi (2004) build overlapping-generations models of wealth inequality that include both intergenerational correlation in human capital and bequests, but neither attempts to model the processes underpinning the correlation in earnings across generations. Bolt et al. (2021) use the same data as in this paper, exploiting the measurement of investments at the youngest ages and outcomes through the oldest working ages for the same individuals to implement a mediation analysis. That paper shows that the mechanisms we consider in this paper are the key ones for explaining the persistence of income across generations. However, they do not estimate preference parameters or model behavioral responses, and so cannot consider counterfactuals.

Since our main policy experiment investigates the impact of student loans, our paper also contributes to the literature on the effects of the financing of education. Reviews of the literature on student loans are provided by Lochner and Monge-Naranjo (2016) and Yannelis and Tracey (2022). The papers closest to ours in that literature are those that study the normative impact of reforms to education financing using models in which altruistic parents make investment decisions on behalf of their children. Abbott et al. (2019) is closely related. They focus on the interaction between parental investments, the financing of education and educational outcomes, but abstract from the role of parents in influencing skill prior to the age of 16. Colas et al. (2021), Krueger et al. (2025) and Fu et al. (2024) also study the interplay between parental investments, state subsidies, educational outcomes, and welfare. Our contributions to this literature are first, in the estimation of the production functions and parental preference parameters using micro data following a cohort for several decades, and in distinguishing between the welfare effects of a reform on different generations who make up the dynasty.

The rest of this paper proceeds as follows. Section 2 describes the data and documents descriptive statistics on skill, education, and the different types of parental transfers. Section 3 lays out the dynastic model used in the paper. Section 4 outlines our two step estimation approach. Section 5 then presents results from the first step estimation, while Section 6 presents identification arguments and results from the second step estimation. Section 7 presents results from counterfactuals and Section 8 concludes.

2 Data and Descriptive Statistics

The key data source for this paper is the National Child Development Study (NCDS). The NCDS started by surveying all children who were born in Britain 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.[3] During childhood, the data include measures of cognitive skills and parental time investments as well as parental income. Later waves of the study record educational outcomes, demographic characteristics,

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

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. Each time we use the NCDS, we use the largest sample for which all required variables are observed. Details can be found in Appendix C. We complement the NCDS with two other datasets paying careful attention that the data is as representative of the NCDS cohort as possible.

First, as the NCDS currently does not have data on the inheritances received or expected, we supplement it using data drawn from similar birth cohorts in the English Longitudinal Study of Ageing (ELSA), which is a representative sample of the 50-plus population in England.

The 2012-13 wave of ELSA recorded lifetime histories of gift and inheritance receipt. We use data on sample members who are born in the 1950s, which gives us a sample of 3,001. Second, we convert the childhood investment measures observed into hours of time using the UK Time Use Survey (UKTUS), which has detailed measures of time spent in educational investments in the child.[4] From this time use data we consider the sum of father's and mother's time spent on the following activities as time investing in children: teaching the child, reading/playing/talking with child, travel escorting to/from education. We provide further details on the data and sample selection in Appendix C.

The rest of this section documents inequalities in parental transfers (time investments, educational investments, and cash transfers) and subsequent outcomes (skill and lifetime income). Throughout the paper we use low, medium, and high to describe education groups – these correspond to having only compulsory levels of education, having some post-compulsory education and having at least some college education, respectively.[5] In the US context, this would correspond roughly to high school dropout, high school graduate, and some college.

2.1 Transfer Type 1: Parental Time Investments

The NCDS has detailed measures of parental time investments received during childhood. The full set of measures we use are listed in Table 1.[6] These measures come from different sources – some are from surveys of parents, others from surveys of teachers. Table 2 highlights some of the key features in the data, conditional on father's education. Since for the parents of the NCDS cohort there is less variation

[4] We use the 1984-1987 waves of UKTUS. These surveys are from slightly later than the NCDS childhood waves. Although a 1961 UKTUS exists, its small sample size limits precision. Nevertheless, the 1961 survey shows a similar level of time investment: 1.5 hours per week on reading, conversation, and games with children, compared to 1.25 hours in the 1984-1987 surveys. Thus, the 1980s surveys balance sample size with accounting for long-term trends in parental time investment.

[5] 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.

[6] While some of these measures are potentially costly in terms of money as well as time, we focus on the time cost, which the previous literature has found to be the key determinant of child cognition (e.g., Del Boca et al. (2014) and Caucutt et al. (2020) for US data, and Bolt et al. (2021) for NCDS data).

in maternal education, we show gradients conditional on father's education, although the gradients by maternal education are similar.

The first panel of Table 2 documents paternal education gradients for some of the investment measures that we use. While 52% of high educated fathers read to their age 7 child each week, only 33% of low educated fathers do so. The gradient is even more pronounced for the teacher's assessment of the parents' interest in the child's education: when the child is 7, 66% of high educated fathers are judged by the child's teacher to be 'very interested' in their child's education but only 20% of low education fathers are. While mothers are assessed as having greater interest in their child's education than fathers, there are large differences according to education group (75% of the highest education group are 'very interested' in their child's education, compared to 33% in the lowest education group).

Table 1: List of all measures used

Skill measures	*Investment measures*
Age 0:	
Birthweight	
Gestation	
Age 7:	
Reading score	Teacher's assessment of parents' interest in education (mother and father)
Math score	Outings with child (mother and father)
Drawing score	Read to child (mother and father)
Copying design score	Father's involvement in upbringing
	Parental involvement in child's schooling
Age 11:	
Reading score	Teacher's assessment of parents' interest in education (mother and father)
Math score	Outings with child (mother and father)
Copying design score	Father's involvement in upbringing
	Parents' ambitions regarding child's educational attainment (further educ & university)
	Parental involvement in child's schooling
	Library membership of parents
Age 16:	
Reading score	Teacher's assessment of parents' interest in education (mother and father)
Math score	Involvement of parents in child's schooling
	Parents' ambitions regarding child's educational attainment

Notes: Age refers to the age when measurements occurred. In our model, we assume investment measures are retrospective, so age investments measured at age 7 are assumed to refer to age 0-6, investments measured at age 11 refer to age 7-10, investments measured at age 16 refer to age 11-15.

2.2 Transfer Type 2: Educational Investments

The second panel of Table 2 shows that there is a substantial intergenerational correlation in educational attainment between fathers and their children. Having a high educated father makes it much more likely that a child will end up with high education. 46% of the children of high educated fathers also end up

Table 2: Transfers and outcomes by father's education

	Avg	SD	Father's education Low	Medium	High	p-val*
Transfer 1: Parental Investments						
Mother reads each week 7	0.49	0.50	0.46	0.56	0.67	0.00
Father reads each week 7	0.36	0.48	0.33	0.44	0.52	0.00
Mother outings most weeks 11	0.54	0.50	0.53	0.61	0.59	0.00
Father outings most weeks 11	0.51	0.50	0.50	0.58	0.56	0.00
Father very interested in education 7	0.26	0.44	0.20	0.43	0.66	0.00
Mother very interested in education 7	0.39	0.49	0.33	0.58	0.75	0.00
Father very interested in education 11	0.31	0.46	0.23	0.52	0.73	0.00
Mother very interested in education 11	0.39	0.49	0.33	0.59	0.76	0.00
Father very interested in education 16	0.36	0.48	0.28	0.57	0.80	0.00
Mother very interested in education 16	0.38	0.49	0.32	0.59	0.78	0.00
Transfer 2: Child Education						
Fraction low education	0.25	0.43	0.30	0.10	0.02	0.00
Fraction high education	0.16	0.37	0.13	0.31	0.46	0.00
Transfer 3: Cash Transfers						
Inter-vivos transfers (>£1000)	0.07	0.26	0.06	0.10	0.20	0.06
Gift value (among recipients only)	39,400	104,600	30,600	77,900	49,100	0.72
Fraction receiving inheritance	0.39	0.49	0.36	0.58	0.54	0.00
Inheritance value (among recipients)	88,200	114,700	75,600	122,400	174,300	0.00
Outcome 1: Child Skills						
Reading 7	0.00	1.00	-0.09	0.33	0.58	0.00
Reading 11	0.00	1.00	-0.13	0.46	0.90	0.00
Reading 16	0.00	1.00	-0.11	0.47	0.77	0.00
Maths 7	0.00	1.00	-0.08	0.26	0.54	0.00
Maths 11	0.00	1.00	-0.13	0.48	0.91	0.00
Maths 16	0.00	1.00	-0.14	0.48	0.99	0.00
Outcome 2: Child's Lifetime Earnings, in £1,000**						
Men	1,347	352	1,289	1,533	1,740	0.00
Women	925	239	879	1,048	1,197	0.00

Notes: For different types of transfers and outcomes, Table 2 shows: Mean, standard deviation, mean conditional on each paternal education group (low, medium, high). 75% of fathers are low education, 20% are middle education, and 5% are high education. Cash transfer data are from ELSA, all other data are from the NCDS. * *P-values* for an *F-test* of the difference in the mean between the low and high father's education group. ** Undiscounted lifetime individual earnings by gender of child, in 2014 pounds, ages 23-55.

with high education, compared to only 13% of those whose fathers have low education.

2.3 Transfer 3: Inter-vivos Transfers and Bequests

The third panel of Table 2 documents the receipt of inter-vivos transfers and bequests, as reported in ELSA, by father's education. The table shows significant differences in the receipt of inter-vivos transfers depending on parental education. Only 6% of individuals from low education families report having received a transfer worth more than £1,000, compared to 20% from high educated families. Moreover, conditional on receipt of a gift, the average value for the two groups differs by about £18,400.

Differences in inheritance receipt by parental background are also significant.[7] 54% of those with high educated fathers have received an inheritance compared to 36% of those with low-educated fathers. Among those who have received an inheritance, those with high educated fathers received more than twice as much on average (£174,300 compared to £75,600). The net result is that those with high educated fathers inherit £66,000 more, on average, than those with low-educated fathers.

2.4 Outcome 1: Skill

The fourth panel of Table 2 shows the average reading and math scores of children at ages 7, 11, and 16, by father's education. As one might expect, those with high education fathers have higher skill levels; at the age of 7, the reading score of children of low educated fathers is 0.09 standard deviations below average, whereas it is 0.58 above average for children of high educated fathers. This gap in reading scores widens with age: by age 16, reading scores of children of low educated fathers is 0.11 standard deviations below average, whereas they are 0.77 above average for children of high educated fathers. Similar patterns are found for math scores.

2.5 Outcome 2: Lifetime Earnings

Finally, we can see that children of more educated fathers have higher lifetime earnings. Lifetime earnings of men with high educated fathers is £1,740k versus £1,289k for those with low educated fathers, a difference of £451k. For women, the difference is £318k.

To summarize, we find that children of more highly educated fathers tend to receive more of each of the three kinds of transfers, and they end up with higher skills, as well as lifetime income. In the following, we present a model bringing together these different types of transfers to explain how these operate in generating the intergenerational persistence in outcomes that we observe.

[7] Sample statistics are calculated for those who lost both parents when interviewed, which is 75% of our ELSA sample.

3 Model

This section describes our dynastic model of consumption, labor supply, and transfers to their children. Figure 1 illustrates the model's timeline. Parental time investments affect children's cognitive skills (which we refer to as skills below). Those skills and educational attainment impact children's wages when they become adults. Upon reaching age 23, children are matched in couples, possibly receive cash transfers from their parents, and begin adult life. They then have their own children and choose consumption, labor supply, and how much to invest in their own children, with implications for their children's future outcomes.

The NCDS interviews respondents every four to seven years from the age of 0 to 55. To be consistent with the data, each of our model periods will cover the time between interviews (and each period will be of different length). Each individual has a life cycle of 20 model periods, which can be broken down into four phases.

1. The Childhood phase has periods $t = 1, 2, 3, 4$ which corresponds to ages 0-6, 7-10, 11-15, 16-22. During childhood, the individual accumulates human capital and education but make no decisions.

2. The Young Adult phase consists of one period at $t = 5$ corresponding to ages 23-25. The individual receives a parental cash transfer (which is potentially 0), is matched into a couple and begins making labor supply and savings decisions.

3. The Parenthood phase has five periods $t = 6, 7, 8, 9, 10$, corresponding to ages 26-32, 33-36, 37-41, 42-48 and 49-54. The couple has identical twin children at the start of the 'Parenthood' phase. In addition to making labor supply and savings decisions, the couple decides how much to invest in their childrens' human capital and education. At the end of this period, the couple can transfer wealth to their children who, in turn, are matched into couples.

4. The Late Adult phase consists of 10 regularly-spaced periods corresponding to ages 55-59, ..., 100-104. The household separates from their children and makes their own saving and consumption decisions.

In outlining the dynastic model, we describe below a life cycle decision problem of a single generation. All generations are, of course, linked. Each couple has children and these children, in turn, will form couples who have children, too. To index generations, we use t to denote the age (in model periods) of the generation of parents whose timeline we outline and a prime to denote their childrens' variables. For example, in the model period when adults are aged t, their children are aged t'.[8]

[8] Children are born five model periods after their parents, therefore they are aged $t' = 1$ in model periods when the parent is model-aged $t = 6$.

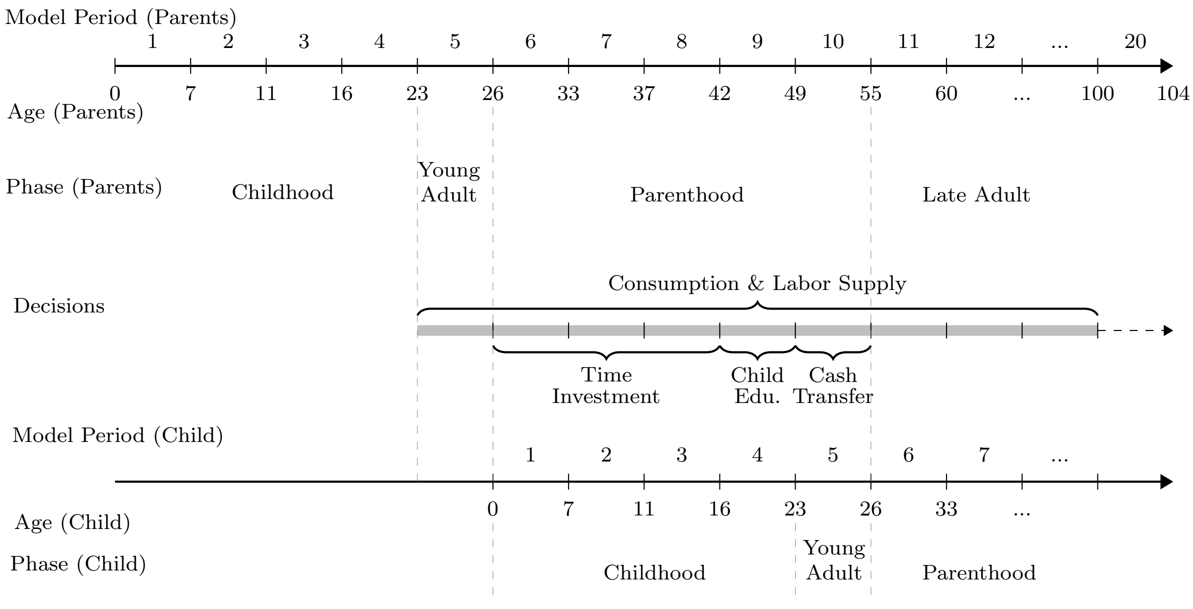


Figure 1: The life cycle of an individual

We now provide formal details of the model.

3.1 Preferences

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

$$u_g(c_{g,t}, l_{g,t}) = \frac{(c_{g,t}^{\nu_g} l_{g,t}^{(1-\nu_g)})^{1-\gamma}}{1-\gamma}$$

We allow preferences for consumption and leisure to vary with gender. Households equally weigh the sum of male and female utility. The household utility function is multiplied by a factor n_t which represents the number of equivalized adults in a household in time t (scaled so that for a childless couple $n_t = 1$).

$$u(c_{m,t}, c_{f,t}, l_{m,t}, l_{f,t}, n_t) = n_t \Big(u_m(c_{m,t}, l_{m,t}) + u_f(c_{f,t}, l_{f,t}) \Big)$$

Total household consumption is split between children, who receive a fraction $\frac{n_t - 1}{n_t}$, and adults, who get a share $\frac{1}{n_t}$. The quantity of leisure is:

$$l_{g,t} = T - (\theta ti_{g,t} + hrs_{g,t}) \tag{1}$$

where T is a time endowment, $ti_{g,t}$ is hours of time investment in children, $hrs_{g,t}$ is work hours, and $l_{g,t}$ is leisure time. $1 - \theta$ is the share of time with the child that represents leisure to the parent: if θ =0 then time with children is pure leisure for the parent, whereas if $\theta = 1$ then time with children generates no leisure value.

The annual discount factor is β. The model period length aligns with the differences in time between interviews and so the discount factor between model periods varies. Thus, the discount rate between t and $t+1$ is $\beta_{t+1} = \beta^{\tau_t}$, where τ_t is the length of model period t.[9]

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

3.2 Demographics

At age 23 all individuals are matched probabilistically into couples, conditional on education. The probability that a man of education ed_m gets married to a woman with education ed_f is given by $Q_m(ed_m, ed_f)$. The matching probabilities for females are $Q_f(ed_f, ed_m)$. The draw of spousal skills and initial wealth is therefore drawn from a distribution that depends on one's own education.

At age 26, a pair of identical twins is born to the couple. In order to match the average fertility for this sample, which is close to two, yet still maintain computational tractability, we follow Abbott et al. (2019) and assume that the twins are faced with identical sequences of shocks.

Mortality is stochastic - the probability of survival of a couple (we assume that both members of a couple die in the same year) to period $t+1$ conditional on survival to period t is given by s_{t+1}. We assume households face mortality risk after the age of 50 and that death occurs by the age of 105 at the latest.

3.3 Human Capital

This section describes the production function for skill and education from birth to age 23. During this part of the life cycle, parental time and educational investments do not directly impact the contemporaneous utility of their children, but leads (in expectation) to the children having higher wages, more highly-skilled spouses, and more highly-skilled childrens' children, all of which matters to the altruistic parent.

3.3.1 Child Skill Production Function

Between birth and age 16, children's skill updates each period according to the production function:

$$\ln(h'_{t'+1}) = \gamma_{1,t'} \ln(h'_{t'}) + \gamma_{2,t'} \ln(ti_{t'}) + \gamma_{3,t'} \ln(ti_{t'}) \cdot \ln(h'_{t'}) + \gamma_{4,t'} ed_m + \gamma_{5,t'} ed_f + u'_{h,t'} \quad (2)$$

[9] In addition, to account for varying period length and within period discounting, we weigh each period's utility by $\sum_{q=0}^{\tau_t} \beta^q = \frac{1-\beta^{\tau_t+1}}{1-\beta}$.

where $h'_{t'}$ represents children's skill when the children are age t'.[10] Children's skill depends on their parents' level of education, the sum of the time investments ($ti_{t'} = ti_{m,t'} + ti_{f,t'}$) that those parents make, past skill, and a shock ($u'_{h,t+1}$). Skill evolves until period 4 (age of 16), after which it does not change.

We allow education of the parents, ed_m and ed_f, to impact skill to capture the idea that high skill individuals who are productive in the labor market may also be productive at producing skills in their children. This is a mechanism that features prominently in several recent studies of child human capital development (e.g., Lee and Seshadri (2019)).

Children's initial skill at birth $h'_{t'=1}$ is a function of their parents' level of education and a shock:

$$\ln(h'_{t'=1}) = \gamma_{4,0} ed_m + \gamma_{5,0} ed_f + u'_{h,0}. \tag{3}$$

3.3.2 Education

When children are age 16, parents choose the education level of their children. There was compulsory education to age 16 for our sample members. Thus, we model the education decision as a choice between sending children to school until age 16, until age 18 (completing secondary education), or until 21 (completing undergraduate education). Because there were no tuition fees for the cohort that we study, we model the cost of education as forgone labor income when at school.

3.3.3 Wages

The wage rate evolves is a function of age, whether the individual works part-time or full-time (where where PT_t is a dummy for working part-time), and an individual specific component v_t:

$$\ln\ w_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + v_t \tag{4}$$

To capture the impact of skill on lifetime wages, we model the initial draw of v_t (in period 5, or age 23) as a function of final skill (h) and a shock: subsequent values follow a random walk

$$v_t = \begin{cases} \delta_5 \ln h + \eta_t, & \eta_t \sim N(0, \sigma^2_{\eta 5}) \quad \text{if } t = 5 \\ v_{t-1} + \eta_t, & \eta_t \sim N(0, \sigma^2_{\eta}) \quad \text{if } t > 5 \end{cases}. \tag{5}$$

Skill impacts the initial wage draw v_5 and so impacts wages at all ages because v_t is modeled as having a unit root. Thus, we do not need to keep track of skill after turning age 23, but instead we keep track of wages as a state variable, which includes v_t and thus final skill. While the associated subscripts

[10] $ti_{t'}$ represents investments the children *receive* when they are aged t', which are equivalent to the investments parents *make* when those parents are aged t.

are suppressed above, each of $\{\delta_0, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \sigma^2_{\eta_5}, \sigma^2_{\eta}\}$ varies by gender (g) and education (ed). This flexibility means that we allow skill to impact wages through its relationship with education δ_5. As we show below, this flexibility is important, as the returns to skill are higher for the highly educated.

3.4 Budget Constraints

Constraints Households face an intertemporal budget constraint, a lifecycle borrowing constraint (no borowing against own future income), and an intergenerational borrowing constrinat (no borrowing against children's future income):

$$a_{t+1} = (1 + r_t)(a_t + y_t - (c_{m,t} + c_{f,t}) - x_t) \tag{6}$$

$$a_{t+1} \geq 0 \tag{7}$$

$$x_t \geq 0 \tag{8}$$

where a_t is household wealth, y_t is household income, and x_t is a cash transfer to children that can only be made when the members of the couple are 49 and their children are 23 (and so $x_t = 0$ in all other periods). The gross interest rate $(1 + r_t)$ is equal to $(1 + r)^{\tau_t}$ where r is an annual interest rate and τ_t is the length in years of model period t.

Earnings and household income Earnings are equal to hours worked (hrs) multiplied by the wage rate; for example: $e_{f,t} = hrs_{f,t} w_{f,t}$. Household net-of-tax income is

$$y_t = \tau(e_{m,t}, e_{f,t}, e'_t, t) \tag{9}$$

where $\tau(.)$ is a function which returns net-of-tax income and $e_{m,t}$ and $e_{f,t}$ are male and female earnings respectively. Before children turn 16, their earnings (e'_t) are 0. Upon turning age 16, children work full time at the median wage given their age, gender, and skill for the part of the period they are not in school. Their parents are still the decision-maker in this period and any income the children earn is part of the parental household income.

3.5 Decision Problem

3.5.1 Young Adult Phase Decision Problem

An individual becomes an active decision-maker at age 23, when they are already formed into a household as part of a childless couple. As such, $t = 5$ is the first model period with a decision problem to solve.

Choices During this phase, couples choose consumption $(c_{m,t}, c_{f,t})$ and hours of work of each adult $(hrs_{m,t}, hrs_{f,t})$ where $hrs_{g,t} \in \{0, 20, 40, 50\}$ hours per week. The resulting vector of decision variables is $\mathbf{d_t} = (c_{m,t}, c_{f,t}, hrs_{m,t}, hrs_{f,t})$.

Uncertainty Couples face uncertainty over the innovation to each of their wages next period $\{\eta_{m,t}, \eta_{f,t}\}$ and the gender and initial skill level of their future children $\{u'_{h,t'=1}, g'\}$.

State variables The state variables $(\mathbf{X}_t)$ during young adulthood are $\mathbf{X}_t = \{t, a_t, w_{m,t}, w_{f,t}, ed_m, ed_f\}$ where t is age; a_t is assets; $w_{g,t}, ed_{g,t}$ are the wages and education of each adult for $\{g \in m, f\}$.

Value function The value function for the young adult phase is given below in expression (10):

$$V_t(\mathbf{X}_t) = \max_{\mathbf{d_t}} \Big\{ u(c_{m,t}, c_{f,t}, l_{m,t}, l_{f,t}, n_t) + \beta_{t+1}\mathbb{E}_t\big[V_{t+1}(\mathbf{X}_{t+1})\big] \Big\} \tag{10}$$

subject to the intertemporal budget constraint in Equation (6) and the borrowing constraints in Equations (7)-(8), where $l_{m,t}, l_{f,t}$ are defined in equation (1) and where the expectation operator is over the innovation to the wage of each of spouse $(\eta_{m,t}, \eta_{f,t})$ and the initial gender and skill of the child $\{u'_{h,t'=1}, g'\}$.

3.5.2 Parenthood Phase Decision Problem: Before Children Reach Young Adulthood

Choices Households choose consumption and hours of work of each adult $(c_{m,t},\ c_{f,t}, hrs_{m,t},\ hrs_{f,t})$. They also choose time investments in children of each adult ($ti_{m,t}$ and $ti_{f,t}$) until their children turn 16, and children's education ed' in the period the children turn 16. The resulting vector of decision variables is $\mathbf{d_t} = (c_{m,t}, c_{f,t}, hrs_{m,t}, hrs_{f,t}, ti_{m,t}, ti_{f,t}, ed')$.

Uncertainty Households face uncertainty over the innovation to each adults' wages $\{\eta_{m,t}, \eta_{f,t}\}$ and the innovations to the childrens' skills $(u'_{h,t})$.

State variables The state variables are $\mathbf{X}_t = \{t, a_t, w_{m,t}, w_{f,t}, ed_m, ed_f, g', h'_t\}$, which is the same as in the independent adult phase plus the childrens' gender (g') and their skill level (h'_t).

Value function The household's value function and the constraints are the same as in equation (10), but with the sets of choices, uncertainty, and states described immediately above.

3.5.3 Parenthood Phase Decision Problem: When Children Become Young Adults

The last period in which parental choices can affect their children is when parents are aged 49 and their children are aged 23 and in the "Young Adult" phase. At the start of the period, parents choose the amount of assets to transfer to their children. Upon receiving this transfer, the children realize their wage draw, get matched into a couple, and start making their own independent decisions.

Choices During this phase, couples choose consumption $(c_{m,t}, c_{f,t})$, hours of work for each parent $(hrs_{m,t}, hrs_{f,t})$, and a cash gift (x_t) which is split equally between their two children and each child recieves a transfer from their in-laws which is a determenstic function of their partners education. The resulting vector of decision variables is $\mathbf{d_t} = (c_{m,t}, c_{f,t}, hrs_{m,t}, hrs_{f,t}, x_t)$.

Uncertainty Couples face two distinct sources of uncertainty. The first is uncertainty over the childrens' initial wage draw and the attributes of their future spouse (his/her skill, education level, assets, and initial wage draw). The second is uncertainty with respect to their own next period wage draws.

State variables The state variables in this phase are the same as in the parenthood phase plus childrens' education (ed').

Value function The value function in this final period of parenthood is:

$$V_t(\mathbf{X}_t) = \max_{\mathbf{d_t}} \left\{ u(c_{m,t}, c_{f,t}, l_{m,t}, l_{f,t}, n_t) + 2\lambda \mathbb{E}_t[V'_{t'}(\mathbf{X'}_{t'})] + \beta_{t+1}\mathbb{E}_t[V_{t+1}(\mathbf{X}_{t+1})] \right\} \tag{11}$$

subject to equations (6)-(8). There are two continuation value functions here. The first is the value function of the (soon to be independent) children. The altruistic parents take this into account in making their decisions. This continuation value function is discounted by the altruism parameter (λ) and the expectation operator is over the children's initial wage draw and the characteristics of their spouse (which are realized after the parents make their decisions). We have assumed that parents have two identical children and make identical decisions about these children and therefore we multiply this continuation value by 2. The second continuation value function is from the expected future utility of the parents when they will enter the late adult phase. This expectation operator is with respect to next period's wage draws and the continuation value is discounted by β_{t+1}.

3.5.4 Late Adult Phase Decision Problem

At this stage, the children have entered their own parenthood phase and the parent couple enters a late adult phase.

Choices Households make labor supply and consumption/saving decisions only ($\mathbf{d_t} = (c_{m,t}, c_{f,t}, hrs_{m,t}, hrs_{f,t})$).

Uncertainty There is uncertainty over next period's wage draws and survival s_t (we assume both members of the couple die in the same period).

State variables The state variables are $\mathbf{X}_t = \{t, a, w_m, w_f, ed_m, ed_f\}$. The skill level and education of the (now grown-up) children are no longer state variables.

Value function Given the definitions of choices, states, and uncertainty for the late life phase, the value function and the constraints take the same form as for the young adult phase (expression (10)).

$$V_t(\mathbf{X}_t) = \max_{\mathbf{d_t}} \left\{ u(c_{m,t}, c_{f,t}, l_{m,t}, l_{f,t}, n_t) + \beta_{t+1} s_{t+1} \mathbb{E}\big[V_{t+1}(\mathbf{X}_{t+1})\big] \right\}$$

subject to equations (6)-(8).

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 production function, the wage process, marital sorting process, and mortality rates. In addition, we also estimate the initial conditions (of the joint distribution of education, skill level, gender, and parental transfers received at age 23) directly from the data. In the second step, we estimate the remaining parameters using the method of simulated moments (MSM) and correct for selection bias in the wage equation.

4.1 Estimating the Human Capital Production Function

Estimating the latent factor production function Exploiting multiple noisy measures of children's latent skill ($h'_{t'}$) and parental investment ($inv_{t'}$) in our NCDS data, we estimate the childrens' skill production function. Following the recent literature (Agostinelli and Wiswall (2022)), we model latent skill as a function of previous period's latent skill, latent investments, parental education, and a shock:

$$\ln h'_{t'+1} = \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'} \ln inv_{t'} + \alpha_{3,t'} \ln inv_{t'} \cdot \ln h'_{t'} + \alpha_{4,t'} ed^m + \alpha_{5,t'} ed^f + u'_{h,t'} \tag{12}$$

We explicitly account for measurement error in the latent factors using a GMM implementation of the methods in Agostinelli and Wiswall (2022). Following the literature (Cunha and Heckman (2008), Cunha et al. (2010)), we assume independence of measurement errors, allowing us to use all possible combinations of the (noisy) input measures to instrument for one another using a system GMM approach described in Appendix D.

Converting latent investments to time Using multiple noisy measures within a latent factor framework helps address the issue of measurement error. However, equation (12) expresses investments as the latent variable inv, whereas the skill production function in equation (2) expresses investments in terms of hours of time. To address this problem, we map latent investment units ($inv_{t'}$) to time units ($ti_{t'}$, which equals the sum of both parents time investments ($ti_{m,t'} + ti_{f,t'}$)). We assume time investments (which we measure using the UK Time Use Survey) with children impact latent investments according to:

$$\ln inv_{t'} = \kappa_{0,t'} + \kappa_{1,t'} \ln(ti_{m,t'} + ti_{f,t'}) \tag{13}$$

where $\kappa_{1,t'}$ is the hours-to-latent investments conversion parameter which determines the productivity of time investments and $\kappa_{0,t'}$ is a constant that ensures we match mean time investments. We allow $\kappa_{0,t'}$ and $\kappa_{1,t'}$ to vary by age, to reflect that the nature and productivity of parental time investments varies by age. We discuss the estimation and identification of the these parameters in Section 6.

With the parameters $\kappa_{0,t'}$ and $\kappa_{1,t'}$ in hand, we substitute equation (13) into the human capital production function (12), which yields the production function we use in our dynamic programming model, equation (2):

$$\begin{aligned}\ln h'_{t'+1} &= \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'}(\kappa_{0,t'} + \kappa_{1,t'} \ln ti_{t'}) + \alpha_{3,t'}(\kappa_{0,t'} + \kappa_{1,t'} \ln ti_{t'}) \cdot \ln h'_{t'} + \alpha_{4,t'} ed^m + \alpha_{5,t'} ed^f + u'_{h,t'} \\ &= \gamma_{0,t'} + \gamma_{1,t'} \ln h'_{t'} + \gamma_{2,t'} \ln ti_{t'} + \gamma_{3,t'} \ln ti_{t'} \cdot \ln h'_{t'} + \gamma_{4,t'} ed^m + \gamma_{5,t'} ed^f + u'_{h,t'}\end{aligned} \tag{14}$$

where $\gamma_{0,t'} = \alpha_{2,t'}\kappa_{0,t'}$, $\gamma_{1,t'} = (\alpha_{3,t'}\kappa_{0,t'} + \alpha_{1,t'})$, $\gamma_{2,t'} = \alpha_{2,t'}\kappa_{1,t'}$, $\gamma_{3,t'} = \alpha_{3,t'}\kappa_{1,t'}$, $\gamma_{4,t'} = \alpha_{4,t'}$, $\gamma_{5,t'} = \alpha_{5,t'}$.

4.2 Identification and Estimation of the Wage Equation

We estimate the wage equation shown in equations (4) and (5), but allow for i.i.d. measurement error in wages u_t. Using those equations and noting that $v_t = \delta_5 \ln h + \sum_{k=5}^{t} \eta_k$ yields an equation for measured wages w_t^*:

$$\ln w_t^* = \ln w_t + u_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + \delta_5 \ln h + \sum_{k=5}^{t} \eta_k + u_t \tag{15}$$

for each gender and education group. Note that by linking latent skills to wages we give a meaningful scale to latent skills h.

In our procedure, we must address three issues. First, wages are measured with error u_t. Second, the final skill level h is measured with error. Third, we only observe wages for those who choose to work, which is a selected sample.

We can address some problems of selectivity using our panel data. To address the issue of composition bias (the issue of differential labor force entry and exit by lifetime wages), we use a fixed effects estimator. Given our assumption of a unit root in $v_t = \delta_5 \ln h + \sum_{k=5}^{t} \eta_k$, which we estimate to be close to the truth (see Appendix G.1), we can allow v_5 (the first shock to wages) to be correlated with other observables (the probability of being observed working at a given age, or part time status), and estimate the model using fixed effects. In particular, we estimate δ_1, δ_2, δ_3, δ_4, and an individual fixed effect FE using a fixed effects estimator:

$$\ln w_t^* \quad = \quad \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + FE + \xi_t$$

where $FE = \delta_0 + \delta_5 \ln h_4 + \eta_5$ captures the time invariant individual specific factors and $\xi_t = \sum_{k=6}^{t} \eta_k + u_t$ is a residual. We then use a methodology similar to that described in Section 4.1 to estimate δ_5 where we use multiple noisy measures of skills to instrument for each other. We then estimate the variances of the wage shocks $(\sigma^2_{\eta_5}, \sigma^2_{\eta})$ and the variance of the measurement error in the wage equation (σ^2_u) using an error components procedure.

While the above procedure addresses problems of measurement error in skill as well as selection based on permanent differences in productivity, it does not address selection based on wage shocks. We account for this last aspect of selection bias by finding the wage profile that, when fed into our model, generates the same estimated profile (i.e., the same δ parameters from equation (15)) that we estimated in the data. Because the simulated profiles are computed using only the wages of those simulated agents that work, the simulated profiles should be biased for the same reasons they are in the data. We find this bias-adjusted wage profile using the iterative procedure described in French (2005). See Appendix G for details for details of estimation of the wage equation as well as our selection correction.

4.3 Method of Simulated Moments

We estimate the rest of the model's parameters (discount factor, consumption weight for both spouses, risk aversion, altruism weight, share of time with the child that represents leisure to the parent and the hours-to-latent investments conversions):

$$\Delta = (\beta, \nu_f, \nu_m, \gamma, \lambda, \theta, \{\kappa_{1,t'}\}_{\{t'=1,2,3\}})$$

with the method of simulated moments, 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. Note that the parameters $\{\kappa_{0,t'}\}_{\{t'=1,2,3\}}$ $\kappa_{0,t'}$ do not enter our list of estimated parameter because they are constants that ensure we match mean time investments. Thus these parameters are a deterministic function of the estimated $\{\kappa_{1,t'}\}_{\{t'=1,2,3\}}$.

Because we wish to understand the drivers of parental labor supply and time investments, we match employment choices for both spouses and also household time spent with children, by parents' age. Because we wish to understand the drivers of education and money transfers, we also match educational attainment, as well as cash transfers to children when the children are older. Because we wish to understand how households discount the future, we match wealth data. Finally, to understand the relationship between time and latent investments, we match observed hours spent with children and their observed skill level. In particular, the moment conditions that we match are:

1. Employment rates, by age, gender, and education, from the NCDS data (30 moments)
2. Fraction in full time work conditional on being employed, by age, gender, and education, from the NCDS data (30 moments)
3. Mean annual time spent with children, by child's age and parent's gender, from the UKTUS data (6 moments)
4. Mean age at which individuals left full-time education by fathers' education level, from the NCDS data (3 moments)
5. Mean lifetime receipt of inter-vivos transfers, from ELSA (1 moment)
6. Median wealth at 60, from ELSA (1 moment)
7. Mean skill at age 16 by father's education, from the NCDS data (3 moments)

We observe individuals in the NCDS, 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 simulate life cycle histories of shocks to skill level, wages, partnering, and childrens' gender and skills for a large number of artificial individuals.

Each individual is endowed with a gender and a value of the age-23 education, wealth, and partner characteristics drawn from the empirical distribution from the NCDS data. Since we do not observe wages of non workers, we draw the initial stochastic component of wages v_5 is drawn from a parametric distribution estimated on the NCDS data.

Next, we solve the model numerically using value function iteration. More concretely, we solve backwards through time over the life cycle of multiple generations using backwards recursion until we find a fixed point in value functions over generations. Because technology (e.g. the wage process and the human capital production function) is assumed constant across generations, expected behavior conditional on the state variables will be the same for all generations. Hence, our assumptions imply the value functions of all generations are the same (i.e., $V_t(\mathbf{X}) = V_t^{'}(\mathbf{X})$). However, the distribution of state variables can vary between generations. For example, one generation can be richer than the next due to increasing education levels, as for the NCDS cohort and their parents.

Using the calculated decision rules in combination with simulated endowments and shocks, we simulate life cycle profiles of behavior for a large number of artificial households, each composed of a man and woman. We simulate life cycle profiles for assets, work hours and time investments, child's educational choices, and inter-vivos transfers. We use the resulting profiles to construct moment conditions, and evaluate the match using our GMM criterion function. We search over the parameter space for the values that minimize this criterion. Appendix I contains a detailed description of our moment conditions, the weighting matrix in our GMM criterion function, and the asymptotic distribution of our parameter estimates. Appendix H gives details of our computational procedures.

5 First Step Estimation Results

In this section we describe results from our first-step estimation, which we use as inputs for our structural model. We present estimates of the effect of parental investments on children's skill, and how that skill affects adult earnings. This exploits a key advantage of our data – that we measure for the same individuals: their parents' investments, their level of skill, and the value of that skill in the labor market.

5.1 The Determinants of Skill

In Section 2, we documented that the children of high educated parents do better in cognitive tests, and that the skill gaps between children of high and low educated parents grow over time. To assess whether this is due to higher parental investments or higher productivity of these investments, we combine multiple test score and parental investment measures to create measures of skills and investments. We use these measures to estimate a skill production function using the methods described in Section 4.1. We estimate equation (12) for skills at ages 7, 11, and 16. The investments entering the equation are those corresponding to ages 0-6, 7-10, and 11-16, respectively. Thus we estimate age 7 skill as a function of age 0 skill, age 0-6 investments, the interaction of skill and investments, and mother's and father's education.

Table 3: Determinants of skills.

	Age 7	Age 11	Age 16
Lagged Skill	0.154	0.739	0.939
	[0.057, 0.251]	[0.696, 0.834]	[0.918, 0.993]
Investment	0.146	0.097	0.131
	[0.113, 0.171]	[0.079, 0.116]	[0.093, 0.161]
Lagged Skill × Investment	-0.021	0.040	-0.038
	[-0.067, 0.010]	[0.027, 0.068]	[-0.066, -0.009]
Mum: Medium Education	0.448	0.181	0.027
	[0.347, 0.552]	[0.109, 0.235]	[-0.026, 0.075]
Mum: High Education	0.593	0.414	-0.088
	[0.388, 0.776]	[0.292, 0.571]	[-0.242, 0.055]
Dad: Medium Education	0.472	0.262	0.056
	[0.252, 0.611]	[0.179, 0.321]	[0.002, 0.115]
Dad: High Education	0.401	0.460	0.107
	[0.313, 0.495]	[0.290, 0.548]	[0.010, 0.218]
Skill shock: $Var(u'_{h,t'})$	0.031	0.067	0.026

Notes: GMM estimates. 90% Confidence intervals [in brackets] are bootstrapped using 100 replications. For the production function at age 7, we use skill measured at age 7 as a function of skill at age 0, time investments measured at age 7 (and referring to investments at age 0-6). For the production function at age 11, we use skill measured at age 11 as a function of skill at age 7, time investments measured at age 11 (and referring to investments at age 7-10). For the production function at age 16, we use skill measured at age 16 as a function of skill at age 11, time investments measured at age 16 (and referring to investments at age 11-15). Skills and investments are in logs.

Estimates are presented in Table 3 (Appendix E gives estimates of the initial skill draw). To ease interpretation, we normalize our skill and measures to have unit variance in every period. Investments have a significant effect on skill, even after conditioning on background characteristics and initial skill. Evaluated at mean skill, a one standard deviation increase in investments at age 0-6 raises age-7 skill by approximately 0.15 standard deviations; a one standard deviation increase in investments at age 7-10 raises age-11 skill by 0.10 standard deviations; and a one standard deviation increase in investments at age 11 raises age-16 skill by 0.13 standard deviations. Skill levels are very persistent, especially at older ages, implying a high level of self-productivity.

The interaction between skills and investments is negative (though insignificant) at age 7 and 16, but positive (and significant) at age 11. However, at all ages, the the interaction terms are modest in size. For example, for those with age 7 log skill levels one standard deviation below (above) mean, a one standard deviation increase in log investment delivers a 0.097-0.040=0.057 (0.097+0.040=0.137) increase in age 11 log skills.

While the richness of our data allows us to account for measurement error in skills and investments, we do not believe our setting allows for credible exclusion restrictions that would allow us to account for the potential endogeneity of investments. The literature has not yet come to a consensus as to whether potential endogeneity would lead us to over- or understate the returns to investments. Attanasio et al. (2020a) and Attanasio et al. (2020b) find that failure to account for endogeneity leads to an understatement of the returns to investments in all periods, whereas Cunha et al. (2010) find that it leads to an overstatement of the returns for older children[11].

We find that parental education strongly impacts future skills, providing empirical support for a key mechanism for perpetuating inequality across generations. High education parents are effective in producing human capital in their children (as also shown in some of the papers cited in the review article by Heckman and Mosso (2014) and is assumed in Becker et al. (2018) and Lee and Seshadri (2019)) in addition to having more resources to afford college. The high productivity of high education parents means that all else equal, their children will have higher skills. As we show below, skills and educational attainment are highly complementary in the production of wages. The combination of these features of human capital production gives high education parents particularly strong incentives to send their children to higher education.

These results are robust to the inclusion of a number of other covariates into the equation, such as parental age and number of children in the household.

5.2 The Effect of Skills and Education on Wages

Wages are a function of age, part time status, age 16 skills, and a shock. By allowing the impact of age 16 skills on wages to depend on educational attainment, we allow the return to skills to vary with educational attainment.

Using the methods described in Section 4.2, Table 4 shows estimates of this impact (δ_5) for each gender and education group. These estimates show the log-point increase in wages associated with a one standard deviation increase in age-16 skill for each education and gender group. The extent of complementarity is similar to that estimated in Delaney (2019) and Daruich (2022), and is implicit in much of the literature on match quality (e.g., Arcidiacono (2005)) and college preparedness in educational choice (e.g., Blandin

[11]Furthermore, Nicoletti and Tonei (2020) find that parents tend to compensate for low cognitive skills, whereas Aizer and Cunha (2012), for example, find that parents reinforce children's skills.

and Herrington (2022)).

Table 4: Log-point change in wages for a 1 SD increase in log skill, by education level

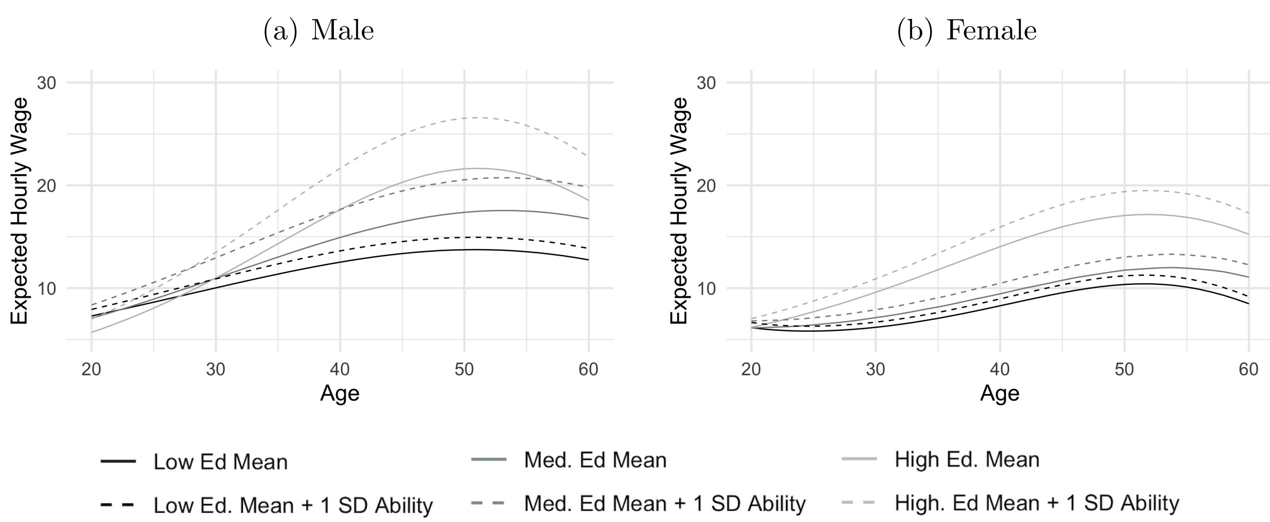
	Male	Female
Low	0.084 (0.025)	0.078 (0.024)
Middle	0.167 (0.019)	0.103 (0.018)
High	0.205 (0.027)	0.127 (0.027)

Notes: Cluster bootstrapped standard errors in parentheses (500 repetitions).

The table shows that, as one would expect, age-16 skill has a significant positive impact on wages conditional on education for all groups. These impacts on wages are larger for for the highly educated. Thus it shows evidence of complementarity between education and skill in the labor market, particularly for men. While low educated men see only a 0.08 log-point increase in hourly wages for every additional standard deviation of skill, high education men (with some college education) see a 0.21 log-point increase in hourly wages for every additional standard deviation of skill. High educated women also receive greater returns to skill than low or middle educated women, although the gradient is more modest compared to that of men.

Figure 2 shows wage profiles by age, education, and gender for full time workers with average skills and also skill levels that are one standard deviation above average, illustrating the complementarity between education and skill.

Figure 2: Wages, by age, education, and gender



Notes: Solid lines are evaluated at mean skill, dashed lines are evaluated at mean skill plus one standard deviation. Lighter lines are for higher education levels, darker lines for lower education levels. Wages measured in 2014 pounds. Wage profiles have been corrected for selection.

As we show below, this dynamic complementarity between skill and education has implications both for optimal time and educational investments. Because of forward looking behavior, households who are more likely to invest in the education of their child have a stronger incentive to invest time in producing skills in their children. Furthermore, those with high skill have an incentive to select into high education.

Turning to the variance of innovations to wages (σ^2_η), Table 5 shows that the estimated variance ranges from 0.0024 to 0.0048, implying that a one standard deviation of an innovation in the wage is 5-7% of wages, depending on the group. These estimates are similar to other papers in the literature (e.g., French (2005), Blundell et al. (2016)). Furthermore, we find evidence that the variance of wage innovations is increasing with education, implying that education is a risky investment.

Interestingly, we estimate the variance of the initial wage shock $\sigma^2_{\eta_5}$ to be small for all groups. While there is significant cross sectional variation in wages, even early in life, we estimate that most of that variation is explainable by our latent skill measure and measurement error in wages.

Table 5: Variance of innovations to wages, by education level

		Men	
	Low	Middle	High
σ^2_η	0.0024	0.0038	0.0045
	(0.0006)	(0.0006)	(0.001)
		Women	
	Low	Middle	High
σ^2_η	0.0020	0.0034	0.0048
	(0.0003)	(0.0004)	(0.0006)

Note: σ^2_η is the variance of the annual innovation to wages. Bootstrapped standard errors in parentheses.

In our formulation, wage shocks have an autocovariance of one: wages are a random walk with drift. This implies skills have a permanent effect on wages. To test this restriction, we also estimated versions of the wage process where we allowed the autocovariance to be less than one. We report in Appendix G.1 little evidence against this restriction and thus use the more parsimonious formulation.

5.3 Marital Matching Probabilities

Table 6 shows the distribution of marriages, conditional on education, that we observe in the NCDS data. It also shows the share of men and women in each educational group. An important incentive for education is that it increases the probability of marrying another high education, and thus, on average, high wage person. Table 6 shows evidence of this assortative matching, as shown by the high share of all matches that are along the diagonal on the table: 12% of all marriages are between couples who are both low educated, 38% are between those who are both middle educated and 4% among those who are both highly educated.

Table 6: Marital matching probabilities, by education

	Low education male	Medium education male	High education male	Share of females in education group
Low education female	0.12	0.19	0.02	0.33
Medium education female	0.13	0.38	0.05	0.56
High education female	0.01	0.07	0.04	0.12
Share of males in education group	0.26	0.64	0.11	

Notes: The numbers represent cell proportions, which are the percentage of all marriages involving a particular match, i.e. these frequencies sum to one. NCDS data, marriages at age 23

5.4 Other Calibrations

Other parameters set outside the model are the interest rate r, parameters of the tax system $\tau(.)$, the household equivalence scale (n_t), time endowment T, and survival probabilities s_t.

The interest rate is set to 4.69%, following Jordà et al. (2019). To model taxes, we use TAXBEN (a tax-benefit microsimulation model developed by the Institute for Fiscal Studies (Waters (2017)) which calculates both taxes and benefits of each family member as a function of their income and other detailed characteristics. We then calculate taxes and benefits (including state pensions) for our sample members at each point in their life, and estimate a three-parameter tax system which varies across three different phases of life: young without children (ages 23-25), working adult (ages 26-64), pension age (age 65 onwards). This three parameter tax system has the following functional form: $y_t = d_{0,t} + d_{1,t}(e_{m,t} + e_{f,t} + e_{f,t} + e'_t)^{d_{2,t}}$.

We set the time endowment T to 5,824 hours per year (16 hours per day). We use the modified OECD equivalence scale and set $n_t = 1.4$ for couples with children. Survival probabilities are calculated using life tables from the Office for National Statistics.

6 Second Step Estimation Results, Identification, and Model Fit

We now present the estimated structural parameters, how they are identified, and the model's fit. Table 7 presents estimates from the structural model.

6.1 Utility Function Estimates and Identification

Table 7: Estimated structural parameters.

Parameter	Estimate
β: discount factor	0.992
	(0.00003)
ν_f: consumption weight, female	0.471
	(0.00007)
ν_m: consumption weight, male	0.464
	(0.00004)
γ: risk aversion	4.14
	(0.006)
λ: altruism parameter	0.267
	(0.0003)
θ: time cost of investment	0.410
	(0.002)
$\kappa_{1,1}$: latent investments per hour, ages 0-6	1.19
	(0.01)
$\kappa_{1,2}$: latent investments per hour, ages 7-10	0.192
	(0.002)
$\kappa_{1,3}$: latent investments per hour, ages 11-15	0.0133
	(0.003)
Coefficient of relative risk aversion, consumption*	2.47

Notes: Standard errors: in parentheses below estimated parameters. β is an annual value. *Average coefficient of relative risk aversion, consumption, averaged over men and women. Calculated as $-(1/2)[(\nu_m(1-\gamma)-1)+(\nu_f(1-\gamma)-1)]$.

6.1.1 Parameters Common to the Literature

Risk aversion. The parameter γ is the coefficient of relative risk aversion (or the inverse of the intertemporal elasticity) for the consumption-leisure aggregate. The coefficient of relative risk aversion for consumption is 2.47 averaging over men and women,[12] which is similar to previous estimates that rely on different methodologies (see Browning et al. (1999) for reviews of the estimates).

Identification of the coefficient of relative risk aversion for consumption is similar to Cagetti (2003) and French (2005) who estimate models of buffer stock savings over the life cycle using asset data as we do. Within this framework, a small estimate of the coefficient of relative risk aversion means that individuals save little given their level of assets and their level of uncertainty. Furthermore, γ is the inverse of the intertemporal elasticity of substitution for utility and thus is key for determining the intertemporal

[12] The coefficient of relative risk aversion for consumption can be obtained using the formula $-\frac{(\partial^2 u_t/\partial c_{g,t}^2)c_{g,t}}{(\partial u_t/\partial c_{g,t})} = -(\nu_g(1-\gamma)-1)$, and so the average is $-(1/2)[(\nu_m(1-\gamma)-1)+(\nu_f(1-\gamma)-1)]$. Note that this variable is measured holding labor supply fixed.

elasticity of labor supply.[13] Life cycle variability in hours and wages provides additional identification, as lower wages early and later in life cause households to substitute from work both into both leisure and time spent with children during these periods.

Discount rate. Our estimate of the time discount factor β is equal to 0.992, and is identified using data on wealth and on labor supply over the life cycle. Both suggest households are relatively patient. First, wealth holdings at age 60 are high given the high level of pension benefits. Second, young individuals work many hours even though their wage, on average, is low. This is equivalent to stating that young people buy relatively little leisure, even though the price of leisure (their wage) is low.

Consumption weight. The parameters ν_m and ν_f are identified by the share of total non-childcare hours devoted to time worked in the market. To see this, note that if hours choice was continuous, the after tax wage would be approximately linked to marginal rate of substitution between consumption and leisure as follows:

$$\begin{aligned} w_{g,t}(1-\tau'_{g,t}) &\leq -\frac{\partial u_t}{\partial hrs_{g,t}} / \frac{\partial u}{\partial c_g} \\ &\leq -\frac{1-\nu_{g,t}}{\nu_{g,t}} / \frac{c_{g,t}}{l_{g,t}} \end{aligned} \tag{16}$$

which holds with equality when work hours are positive, where $\tau'_{g,t}$ is individual g's marginal tax rate at time t.[14] Inserting the time endowment equation (1) into equation (16) and making the approximation that consumption equals after tax earnings $c_{g,t} \approx w_{g,t} hrs_{g,t}(1-\tau'_{g,t})$ yields

$$\nu_g \approx \frac{hrs_{g,t}}{T - ti_{g,t}}. \tag{17}$$

Thus ν_g is approximately equal to the share of non-childcare hours that is spent at work. We find that this share is somewhat less than 0.5, and thus our estimate of ν_g is modestly less than 0.5 for both men and women.

Altruism. Our estimate of the weight that the altruistic parents place on the utility of both their children (2λ) is 0.53, which is the middle of the range of estimates reported in the literature. This is higher than estimates by Daruich (2022), who estimated it to be 0.36 and Lee and Seshadri (2019), who estimate it to be 0.32, and lower than Gayle et al. (2022), whose estimate is 0.80 and Caucutt and

[13]Assuming certainty, linear budget sets, and interior conditions, the Frisch elasticity of leisure is $\frac{\nu_g(1-\gamma)-1}{\gamma}$ and the Frisch elasticity of labor supply is $-\frac{l_{g,t}}{hrs_{g,t}} \times \frac{\nu_g(1-\gamma)-1}{\gamma}$. However, an advantage of the dynamic programming approach is that it is not necessary to assume certainty, linear budget sets, or interior conditions.

[14]This relationship is not exact because of the part time penalty to work hours and the discreteness of the hours choice.

Lochner (2020), whose estimate is 0.86. These papers model a parent with only one child, whereas in our framework parents have two children. Thus we multiply by 2 the continuation values of the children.

The parameter λ is identified from two sources. First, households make cash transfers to their children. These are the most direct manifestation of altruism. To see this, note from equation (11) that in the phase when the child is in their young adult phases ($t = 9$, when the parent is 49 and the child is 23), parents have the opportunity to transfer resources, and the following optimality condition holds

$$\frac{\partial u_t}{\partial c_{g,t}} \geq 2\lambda \frac{\partial \mathbb{E}_t V'_{t'}(\mathbf{X'}_{t'})}{\partial A'_{t'}} = 2\lambda \mathbb{E}_t \left(\frac{\partial u'_t}{\partial c'_{g,t'}} \right)$$

and holds with equality if transfers are positive. The term on the right is the sum (over both children) of the childrens' expected marginal utility of consumption. At the time of the transfer, the children will be at a low point of their lifecycle earnings, and will soon have their own children with the accompanying time and money expenses. This, and the fact that they are likely to be borrowing constrained, will mean they will have a higher marginal utility of consumption than their parents. In order to rationalize relatively modest transfers to children, λ must be less than 1. Nevertheless, the fact that these transfers are made is perhaps the strongest evidence that $\lambda > 0$ and households are altruistic.

Second, λ is identified from household investments in the educational attainment of their children. The foregone household income from children going to school represents a direct loss of resources to the household. In particular, the returns to education average 7.4% per year of education, which is well above the market interest rate of 4.7%. Recall that the returns to education accrue to the child, whereas the return to cash accrues to the parents. The fact that many parents invest little in their childrens' education, but some invest a lot, again provides again evidence that λ is less than 1 but is greater than 0.

6.1.2 Novel Parameters

Taking stock of our progress so far, we have provided identification arguments for all model parameters except two: the time cost of investment (θ) and the mapping from time investments to latent investments $\{\kappa_{1,t'}\}_{\{t'=1,2,3\}}$. Our heuristic discussion of identification of the other parameters did not rely on data on parental time investments in children or children's skills. Here we describe heuristically how we use those data to identify these remaining parameters (θ and $\{\kappa_{1,t'}\}_{\{t'=1,2,3\}}$), taking the other parameters as given.

For these remaining parameters, observed time investing in children plays a central role. Central to the identification argument is the following observation: parents choose to spend time with their children either because (i) the effective cost per unit of latent investment is low (i.e., the mapping $\kappa_{1,t'}$ is high, implying high returns to time invested), or (ii) because parents derive direct utility from the time itself (i.e., θ is low). Holding constant observed time investments, the former channel affects the children's skills, while the latter does not as it reflects preferences over time use. This distinction means that, together, observed variation in time inputs and skill outcomes allows us to disentangle these effects and identify both sets of parameters.

Latent investment per hour. The parameters $\kappa_{1,t'}$ govern the relationship between time inputs and latent investments. Their identification rests on the observed relationship between time spent with children (from UKTUS data) and children's final age 16 skills (from NCDS data), and does not depend on any assumptions about household optimization.

In particular, the combination of variation in time spent with children and in children's final skills (by father's education) allows us to discipline $\kappa_{1,t'}$. If these parameters are set too high (low), the model will predict too much (too little) dispersion in child skill outcomes by father's education. Thus, the joint distribution of observed time spent with children and children's skills pins down $\kappa_{1,t'}$. We provide a more formal argument in Appendix K.

Time cost of investment. The parameter θ is identified through the trade-off between time invested in children and other uses of parental time, as implied by household optimization. The key idea is that the observed amount of time spent with children must reflect an optimal choice, balancing the returns to child investment with the opportunity cost of time.

Conceptually, $1 - \theta$ captures the share of time with children that parents perceive as leisure. If $\theta = 1$, time with children is as costly (in utility terms) as market work; if $\theta = 0$, it is as enjoyable as pure leisure. If $\theta = 1$, optimal behavior implies that the economic benefit of an additional hour invested in the child (i.e., the expected increase in lifetime earnings) must roughly equal the parent's wage since the parent's outside option is market work. Thus the parent should be indifferent between spending time with the child or using that time to earn money to transfer to the child. Conversely, when $\theta = 0$, parents would choose to spend time with their children even if it had no return in terms of child outcomes.

Our estimates suggest that parental time with children is neither purely work nor purely leisure: we obtain an estimate of $\theta = 0.41$, implying that 59 percent of the time parents spend with children is perceived as leisure, while 41 percent is viewed as a costly investment. Appendix K provides a formal treatment of identification in the context of a simple model.

There is limited direct empirical evidence on the magnitude of this parameter. Some studies assume that parental investments can be fully explained by the financial returns resulting from increased human capital, while others assume that parents derive some additional utility benefit from time spent investing in children's human capital. But to the best of our knowledge, no other paper estimates this parameter while nesting the possibility that parents treat time with children identically to work.

Caucutt and Lochner (2020) and Lee and Seshadri (2019) are illustrative of studies assuming parental investments are explained by higher future wages of their children. These papers match their models to the data under the assumption that parental investment can be explained by the returns to human capital, each having a parameter (parental altruism and the conversion of child to adult human capital, respectively) that adjusts to make this possible. As they only model a subset of the intergenerational transfer channels, they do not have the overidentifying restrictions that could lead to rejection of this assumption. It is only because we can identify the extent of parental altruism from education and cash transfers that we can tell that parental time with children is higher than implied by the return to parental time.

Intriguingly, studies that assume parents view time with children differently than work tend to model a wider range of channels (e.g., Krueger et al., 2025; Daruich, 2022; Fuchs-Schündeln et al., 2022). However, these studies tend to be more macro-focused, calibrating many key components of their models from the literature and utilizing moments from a wide range of datasets across various cohorts. The more macro approaches taken in these papers, combined with the fact that many do not nest the possibility that parents view time investments in children as identical to work, limit their ability to test this assumption. It can also be difficult to make like-for-like comparisons with the parameter values these papers find, as their approaches are often slightly different from ours. For example, Daruich (2022) models all time spent with children as part of leisure but adds a separate disutility from time spent with children. That said, he also finds a substantial role for parental utility from time with children. When we convert his values into an equivalent θ, we find that this can be as low as 0.02 for full-time workers at average time-use values.

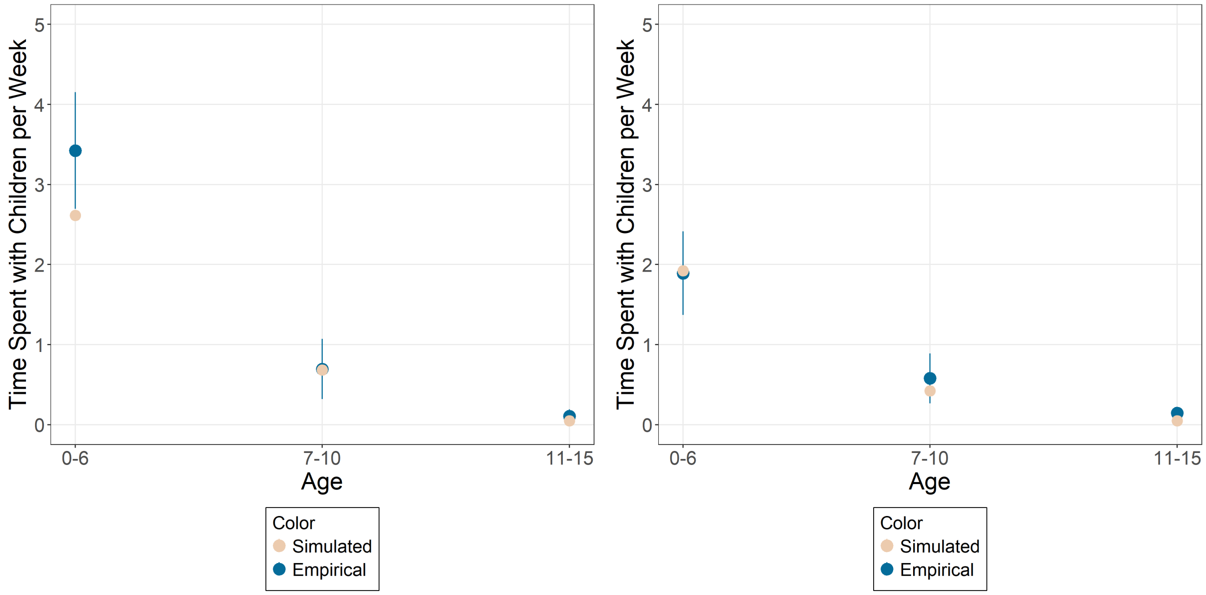
6.2 Model Fit

In this section, we focus on the moments that are critical for understanding intergenerational altruism: transfers of time, educational investments, and money.

Figure 3 shows transfers of time from mothers and fathers in the left and right panels, respectively. The model fits two key patterns in the data well. First, time investments decline with age. Second, mothers invest more in their children than fathers. This higher rate of investment reflects the lower wage, and thus the lower opportunity cost of time for women.

Our model captures the final gradient of skill levels of children with respect to children's education, as shown in Figure 4(a). Children of low education fathers have skill levels that are 0.14 standard deviations below average, whereas children born to high education fathers have skill levels that are 0.80 standard deviations above average. Our model matches these patterns well, although we slightly understate the gradient.

Figure 3: Model fit: parental time with children



(a) Mother's Time with Children

(b) Father's Time with Children

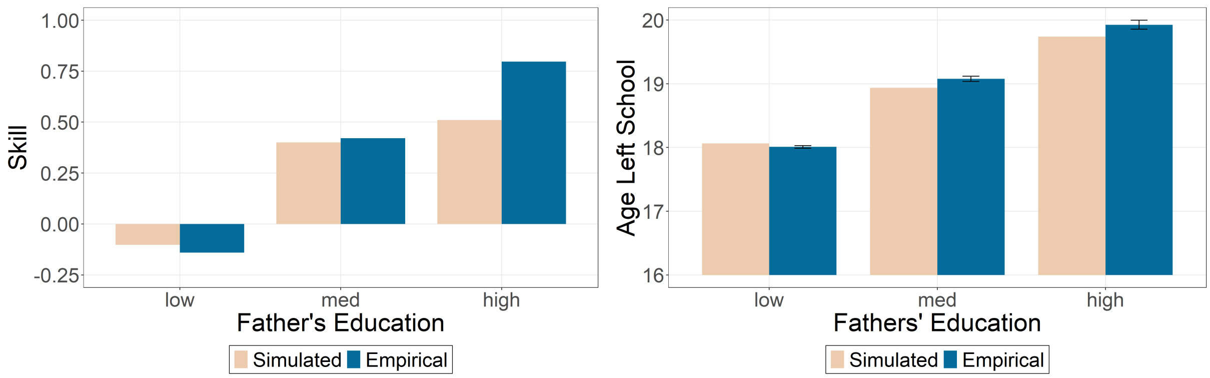
Notes: Parental time investments. Source: UKTUS. See Appendix C.3 for details.

Next, panel (b) of Figure 4 shows children's education by father's education. The model captures the gradient of children's education by parent's education. The difference between the average age left school of the children with high educated fathers and those with low educated fathers that our model predicts is 1.68 years versus a difference of 1.92 years found in the data.

Table 8 shows that we match well the mean level of financial transfers received and the median level of assets at age 60. These financial transfers include inter-vivos transfers when younger and bequests received when older. These amounts are discounted to age 23 (the age at which transfers happen in the model); when undiscounted, the amounts are considerably larger. In the data, as in the model, median transfers are 0. Thus, we match mean transfers.

Finally, our model can reproduce key labor supply moments of men and women with different education levels, as shown in Appendix J. The model does well in generating a dip in female participation and full time work between ages 33 and 48 (when children are in the household). Moreover, as in the data, the model predicts higher participation rates for more educated women at older ages. For men, the model does well in generating a level of labor supply that is consistent with the data both on the intensive and the extensive margin.

Figure 4: Model fit: education and skill



(a) Skill level by father's education (b) Education choices by father's education

Notes: Educational attainment and skills from NCDS data.

Table 8: Model fit: transfers and assets

	Empirical	Simulated
Mean transfers	£12,900	£15,200
Median Assets	£306,400	£246,100

Notes: Values in 2014 GBP. Mean transfers and median assets calculated using ELSA data. Transfers include inter-vivos transfers and bequests and are discounted to age 23 at the real rate of return. See Appendix C for more details.

6.3 Intergenerational Persistence

Although we do not target these directly, our model replicates the intergenerational persistence in economic outcomes that are commonly estimated in other studies. This includes the intergenerational correlation of education and the intergenerational elasticity (IGE) of lifetime earnings and consumption. We estimate the following regression on our simulated data: $y' = a_0 + a_1 y + u$ where y' denotes the child's outcome (for example, the number of years of schooling or the log of childrens' household earnings) and y the parents' corresponding outcome (for example, parents' years of schooling or lifetime household earnings).

The model-predicted correlation of childrens' and parent's education is 0.30 which is similar to the estimates presented in Hertz et al. (2008) who report 0.31 for Great Britain. The model-predicted intergenerational elasticity of lifetime household earnings is 0.50, which is slighltly higher than the estimated values reported in Blanden et al. (2007) and Bolt et al. (2021).

Table 9: Intergenerational Persistence

Outcome	Model-Implied	Literature
Intergenerational Correlation, Education	0.30	Hertz et al. (2007) ≈ 0.31
Intergenerational Elasticity, Earnings	0.50	Blanden et al. (2007), Bolt et al. (2021) ≈0.35
Intergenerational Elasticity, Wealth	0.27	Charles and Hurst (2002) ≈0.37 (in the US)

Notes: Intergenerational correlations and elasticities calculated from model simulated data. Earnings and consumption calculated as average over ages 23-65.

Table 10: θ fixed equal to 1

Outcome	Baseline $\theta = 0.41$	Experiment $\theta = 1.00$
Lifetime wage	454,000	416,000
Age Left Education	18.8	19.0
Fraction high educated	51%	57%
Transfers	£15,200	£16,700
Skill at 16	0.28	-0.43
Intergenerational Correlation Education	0.30	0.27
Intergenerational Elasticity of Earnings	0.50	0.40
Intergenerational Elasticity of Wealth	0.27	0.21

Notes: The left column contains the results from the baseline simulations, serving as a reference. The right column contains simulations of a model in which parents view time with children as a pure investment.

6.4 Sensitivity to utility of time investment (θ)

Since θ is a novel parameter, and the standard assumption in much of the literature is to assume $\theta = 1$, we conduct an experiment to assess the sensitivity of our results to θ. If we make this particular form of investment more costly, we would expect parents to substitute towards other forms of investment (e.g. educational attainment). Thus, the overall impact on earnings is therefore ambiguous.

Column 2 of Table 10 contains the results of this experiment, where we fix $\theta = 1$ so that parents view time with children as being as enjoyable as work. Column 1 contains the baseline results for reference. Educational attainment and cash transfers increase as parents substitute away from time investments, but the lifetime wages of the children decrease due to the corresponding decrease in final skills. The wage decrease is substantial, representing 8.3% of the baseline wage, indicating that the parental preference for time with children plays a significant role in shaping earnings power.

7 Results

7.1 Returns to time with children

We use our model to calculate the lifetime wage returns to time spent with children. This requires (i) the estimated wage returns to skill (via the wage equation), (ii) the estimated skill returns to latent parental investments (through the skill production functions), and (iii) the estimated mapping from hours of parental time to latent investments. Understanding the returns to time spent with children is central to understanding responses to various policy proposals.

It has been well established that early life investments can have a return well in excess of the market return. However, this has been established in the context of professional childcare (e.g., Heckman et al. (2010)). Furthermore, a large literature has documented a clear positive association between measures of the intensity of parental investment (but not time itself) and subsequent skills and wages (e.g., Cunha et al. (2006)). But without hours of time itself, it is impossible to calculate the return to parental time. We bridge the gap and find that the return to parental time invested in children is below the market wage for most parents.

To do this, we re-simulate the model, holding all behaviors constant except for time spent with children during the first period of life (ages 0 to 6). In this period, we increase parental time by one additional hour per week for each of the seven years in this time period, resulting in a total increase of 364 hours. We then compute the increase in the children's lifetime wage relative to that generated by the baseline model, with the value discounted back to the start of the children's life.

Table 11 shows that this additional hour per week of parental time over the first seven years of life raises the each child's lifetime wage by an average of £1,440, or £2,880 when summing over the two children. This corresponds to a return of £3.95 per child, or £7.90 when summing over the two children, per hour of parental time.

These average returns mask substantial heterogeneity. For children who go on to higher education, the return to an hour is on average £11.00, but for those who stop at compulsory education, the return is only £3.70. These differences largely reflect the higher return to skills for those with high educational attainment.

Children of more educated parents experience greater returns to parental time, for two main reasons. First, they are more likely to attend college and thus benefit from the high returns to skill associated with higher education. Second, their parents are more effective at producing skills, as parental education enters directly into the skill production function.

Similarly, children with higher initial skills also gain more, for reasons that parallel those above. First, they are more likely to pursue higher education and therefore benefit more from the complementarity between skill and educational attainment. Second, they are more likely to have highly educated parents who are more productive in investing in their development, making each additional unit of investment more effective in levels.

Table 11: Returns to an hour with children

Outcome	Increase in lifetime wage	Per hour
Mean Increase in child's wage per parental hour	£2,880	£7.90
... *among those who obtain low education*	£1,340	£3.70
... *among those who obtain medium education*	£2,600	£7.10
... *among those who obtain high education*	£4,020	£11.00
... *among children of low education fathers*	£1,730	£4.80
... *among children of medium education fathers*	£3,130	£8.60
... *among children of high education*	£4,010	£11.00
... *among children with below median initial skill*	£2,740	£7.50
... *among children with above median initial skill*	£3,030	£8.30

Notes: Change in children's wage resulting from an additional 364 hours of parental investment, given as one additional hour per week from 0 to 6 years old, while holding all other behaviors constant. Wages are discounted back to age 23 (the age at which the children become independent). Values presented are for the sum of the wage gains over both children that the couple has.

These returns to an hour of parental time investments are largely below the wage rates shown in Figure 2 (and are even below the wage of the lower earning parent who is typically the mother). The difference between the return to time investments and the wage rate help explain our estimate of θ. If parents do not value time with children as partly leisure, it is very difficult to rationalize the observed patterns of parental investment. If the only returns to time investments in children are through the impact on children's future earnings, it would be optimal for almost everyone to work more and transfer money to the child later in life.

7.2 How is Income Risk Resolved over the Life Cycle?

How much of the cross-sectional variance in lifetime income can we predict using information known at different ages? Already before birth, information on the parents can help us predict an individual's lifetime income through predicted future investments that parents will make as well as through the productivity of those investments. As the child is born and grows older, decisions are made, and shocks are realized, thus increasing the extent to which lifetime income can be predicted.

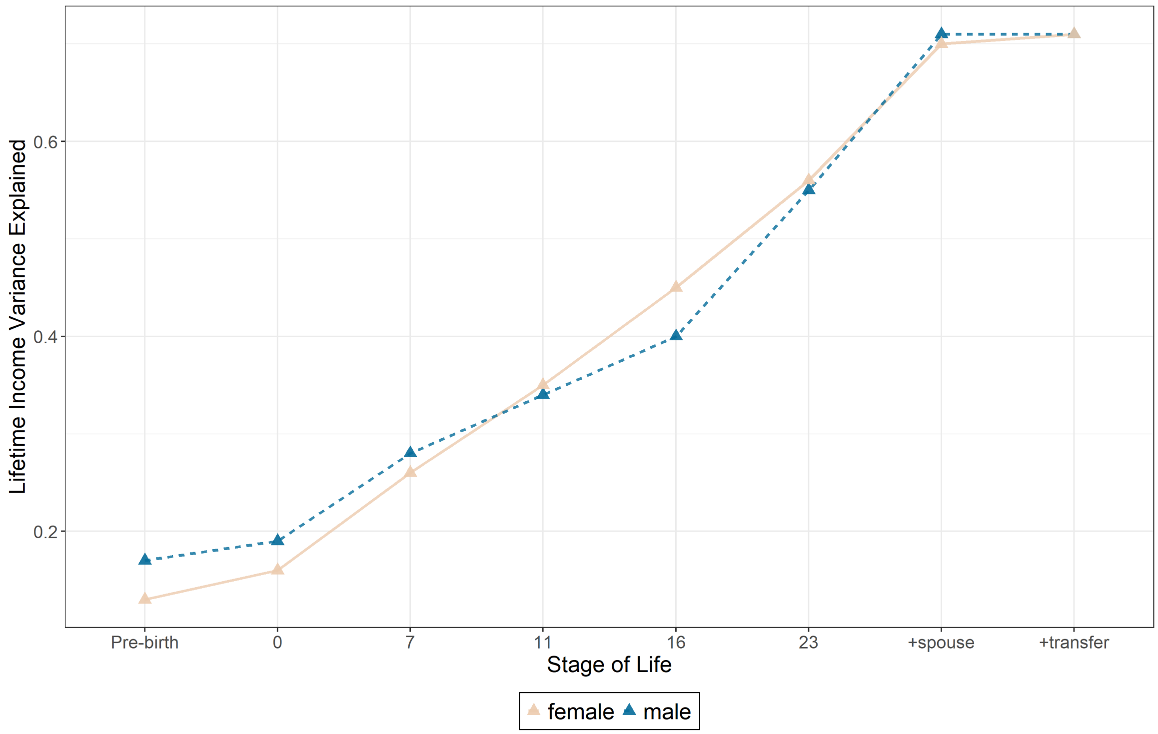
We estimate how uncertainty is resolved for agents starting from before birth. We take as given the age-23 joint distribution of the state variables of their parents. In particular, we take as given the age-23 joint distribution of the state variables of the NCDS sample members, draw histories of shocks, and calculate optimal decisions for both NCDS sample members and their children. This allows us to simulate lifetime outcomes for two generations. Next, we calculate the share of the variance in the childrens' lifetime household income (including transfers) that can be predicted by the following variables that are known at each age: parental assets, wages, and education; the child's skill level, gender, education, and wages; and (once the child is 23) the education and wages of the child's spouse. This approach allows us to decompose the relative importance of (predictable) circumstances and choices; the remainder being explained by shocks. This builds upon the approach in Huggett et al. (2011), who calculate the share of lifetime income known to the individual at age 23, and Lee and Seshadri (2019), who calculate it before birth, at birth, and at age 24. By showing the amount of lifetime income variability known at multiple ages, we illustrate how this uncertainty is resolved with age and how it is resolved after marriage. This decomposition cannot be carried out without the model, as although the NCDS contains data on a single generation from birth to retirement, it does not contain this data for two generations.

Our decomposition makes use of the law of total variance: a random variable can be written as the sum of its conditional mean plus the deviation from its conditional mean. As these two components are orthogonal, the total variance equals the sum of the variance in the conditional mean plus the variance around that mean. We then divide the variance in the conditional mean of lifetime income by the total variance. We illustrate how uncertainty about household income is resolved over the lifecycle in Figure 5.

The figure shows that 17% and 13% of the variance lifetime household income are known for males and females, respectively, even before they are born. These shares are explained by parents' education (which affects initial skill and the productivity of parental investments) and also household financial resources (which affects the quantity of investments the child receives). As the child ages, new information is realized, both about their own skill and their parents' financial resources. Immediately after birth, initial skill and parental wage shocks are revealed, causing the shares explained to rise to 19% and 16% for males and females, respectively. By the time the children are aged 23, educational choices have been made and their initial wage draw has been realized, causing the shares to rise to 55% and 56%. Thus, over half of lifetime wage variability is realized by age 23.

At age 23, individuals marry, resolving uncertainty about the spouse's wage, education, and parental transfer. Marriage explains much of the remaining variability in lifetime household income: the share explained jumps to 70% for both men and women after marriage. Before marriage, the characteristics of one's future spouse is an important risk; after marriage, one's spouse becomes an important form of insurance. Furthermore, transfers explain little of lifetime income, both because transfers are small relative to lifetime earnings and also because the transfers made are highly explainable given all the other variables known.

Figure 5: Resolution of uncertainty over the life cycle



Notes: "male" denotes male household income, "female" for females, analogously. "+spouse" denotes age 23 after being matched into a couple. "+transfers" denotes age 23 after transfers from parents received. This graph shows the share of variance explained by characteristics of both parent and child known at a given age of the child. Income is discounted pre-tax values. Household income is the sum of earnings plus parental transfers received by both the individual and their spouse.

7.3 Relaxing Intergenerational Borrowing Constraints through Student Loans

The financing of university education is central to policy debates in many countries. Student loans are widely used to lower public costs and expand access, but they are often unpopular with young people, highlighting a generational divide in attitudes to student debt (e.g., CNN, 2022). We discuss this intergenerational tension through the lens of our model.

Our model includes an intergenerational borrowing constraint that prevents parents from borrowing against their children's future earnings. We consider a student loan policy that relaxes this constraint, with two conditions reflecting typical features of student loan schemes. First, borrowing is allowed only if the child attends university. Second, borrowing is capped at the opportunity cost of university—namely, the wages forgone by not working during that period. In our model, this results in a cap of approximately £37,000.

We re-solve the model under the student loan policy to assess its impact on lifetime wages, educational attainment, intergenerational transfers, skill formation, and the intergenerational persistence of outcomes. We assume the policy is anticipated from the beginning of the parent generation's working life, allowing

Table 12: Model Experiments

			Policy	
	Outcome	Baseline	Student Loans	Targeted Student Loans
(1)	Lifetime wage	454,000	463,000	467,000
(2)	Age Left Education	18.8	19.5	19.2
(3)	Fraction high educated	51%	69%	61%
(4)	Cash Transfers	£15,200	-£2,100	£5,700
(5)	Skill at 16	0.28	0.23	0.35
(6)	Intergenerational Correlation Education	0.30	0.25	0.34
(7)	Intergenerational Elasticity of Earnings	0.50	0.43	0.40
(8)	Intergenerational Elasticity of Wealth	0.27	0.12	0.16
(9)	Utility change (£ terms)		-£2,000	£1,300
(10)	*...among children high educated in baseline*		-£31,000	-£18,700
(11)	*...among children not high educated in baseline*		£34,000	£21,800
(12)	*...among children of low-ed father*		£1,800	£5,000
(13)	*...among children of medium-ed father*		-£800	£2,300
(14)	*...among children of high-ed father*		-£13,200	-£9,200

Notes: The leftmost column contains the results from the baseline simulations, serving as a reference. The Student Loans column contains results from simulations of a version of the model in which parents can borrow against the children's income up to a limit of £37,000. Lifetime wages, transfers, and the compensating asset measuring the utility change are discounted back to the first period of the working life. Targeted student loans are loans that go only to those in the top half of the ability distribution. Utility change is the compensating asset that makes the individual indifferent between the baseline no student loan case and the student loan case. A positive value means they would be willing to pay for the policy.

parents to adjust their time and educational investments in response to the policy.

In analyzing this reform, we account for the equilibrium effect of the policy on the marriage market. In particular, the subsidy affects not only the education of an individual but also the distribution of educational levels within the economy and, therefore, the distribution of potential spouses. Thus, the marital matching probabilities change. Although we do not impose an equilibrium matching model, our approach respects marriage market clearing by exploiting historical variation in marriage matching probabilities conditional on education as documented in Appendix M.

Table 12 summarizes the results. Relative to the baseline (column 1), student loans (column 2) raise average lifetime earnings by £9,000 (from £454,000 to £463,000 in row 1), driven by increases in education: the average age of school leaving rises by 0.7 years (row 2), and university attendance increases by 18 percentage points (row 3). The patterns are qualitatively consistent with the evidence exploiting discontinuities in loan access as students in Chile cross a GPA boundary in Aguirre (2021) and that in Abbott et al. (2019) who estimate a dynastic model of parental investment.

However, parents also reduce direct transfers (row 4) by over £17,000 on average. Average transfers become negative as parents take student loans that their children must re-pay. Final skill at age 16 falls slightly (row 5), reflecting a shift of family transfers from time investment to educational attainment. Rows 6-8 show declines in the intergenerational persistence of education, earnings, and wealth.

From the perspective of the dynasty, the student loan policy is a pure relaxation of borrowing constraints and therefore expands the choice set. As a result, dynasties would be willing to pay, on average, £13,090 (in present value terms at the start of the child's working life) to access such a policy. However, the policy is not necessarily welfare improving for the children's generation, especially if altruism parameter (λ) is low. To assess the welfare change to the children's generation, we calculate the compensating asset transfer required to make children indifferent between the baseline and student loan cases. On average, this is -£2,000 (row 9), implying that many members of the children's generation would have preferred the policy not be introduced.

These welfare effects vary sharply by whether individuals would have attended college absent the policy. Among infra-marginal individuals—those who would have gone to university regardless—the policy imposes a welfare loss of £31,000 (row 10), as it enables parents to borrow against their future income without altering their educational outcomes. By contrast, marginal individuals who are induced to attend college by the policy experience welfare gains of £34,000 (row 11), reflecting the benefit of increased education and earnings.

This difference in impact on marginal and infra-marginal individuals leads to important differences in impact by socioeconomic status. In the final panel of Table 12, we observe that the policy's impact on the children of fathers with medium and high levels of education is substantial and negative, with a significantly larger negative impact on the children of highly educated fathers. These groups of children are most likely to attend university in the baseline and, therefore, contain a larger share of infra-marginal individuals. The policy does, however, have a positive impact on the children of low-educated fathers, as this group of individuals contains the largest share of marginal individuals who can benefit from the additional opportunities to attend college. This counterfactual highlights the potential for more targeted policies to improve welfare not only for the parents' generation but the children's generation as well.

This student loan experiment is universal in the sense that anyone who chooses to go to college can receive the loan. The third column of Table 12 presents the results of a more targeted merit-based loan scheme under which only those attaining above the median skill level observed in the baseline would be eligible. We observe that this conditionality improves the welfare implications of the policy for the children's generation who, on average, benefit. It also leads to a large increase in the final skill of the children, as parents make additional time investments in order to be eligible for the policy, and average cash transfers become positive as fewer parents are able to borrow against their children's future earnings. However, compared to the unconditional loan policy, this policy leads to a smaller reduction in intergenerational persistence. This is because higher-educated parents are better able to invest in their children's abilities, both because they have more resources and because they are more productive at producing their children's skills.

Together, these results reveal a central tradeoff in the design of student loans. Universal access reduces intergenerational persistence but may impose costs on the very generation it aims to support. Merit-based schemes improve average welfare but disproportionately benefit children of already-advantaged parents, reinforcing existing inequalities. This tension between efficiency and equity, and between immediate welfare and long-term mobility, lies at the heart of debates over how best to structure educational finance.

8 Conclusion

This paper estimates a dynastic model of parental altruism where parents can invest in their children in multiple different ways. We estimate child skill production functions, wage equations, and preference parameters using our model and data from a cohort of children followed from birth, allowing us to measure how parental investments early in life impact wages in adulthood. Our estimates reveal that parents view time with their children as serving a dual purpose: it is both an investment in the child's future and a source of direct enjoyment. As a result, the opportunity cost of time with children lies between that of work and leisure, which implies that time investments are not solely motivated by their financial returns.

We use the estimated model to analyze how intergenerational borrowing constraints shape parental behavior and child outcomes. We show that relaxing intergenerational borrowing constraints can increase educational attainment and reduce intergenerational persistence in earnings and wealth. However, the welfare impact on the children's generation of offering unconditional loans for college depends crucially on what the impacted children would have done in the absence of the availability of the loan. Welfare gains accrue to children who would not have attended college otherwise, while infra-marginal students, who would have gone to college regardless, bear new debt burdens without new benefits.

These findings underscore a core tension in the design of programs which relax intergenerational borrowing constraints. Universally available loans promote mobility but risk imposing costs on the very cohorts they aim to help. Targeted, merit-based loans can improve average welfare, but do so by amplifying the advantages of some already well-off families. More broadly, our results show that understanding the intergenerational transmission of advantage requires a model that captures not just endowments and constraints, but parental preferences over the time that they invest in their children's skills, and how they value their children's realized outcomes.

While our model captures many key channels of intergenerational transmission, we leave many interesting issues for future research. First, we abstract from capacity constraints in higher education:

our simulations assume that all children who benefit from the loan policy are able to attend college, regardless of broader institutional limits. Gu and Zhang (2024) study such constraints. Second, we hold factor prices fixed and do not model general equilibrium responses in wages or returns to education that could arise if college enrollment expands substantially (see for example Heckman et al. (1998) and Abbott et al. (2019)). Finally, we do not incorporate the likely positive effects on government revenues that the additional loan-funded education would generate (see Fu et al. (2024)). Incorporating these forces may dampen or amplify some of the effects we estimate and remains an important direction for future work.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy 127*(6), 2569–2624.

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

Agostinelli, F. and M. Wiswall (2022, July). Estimating the technology of children's skill formation. Working Paper 22442, National Bureau of Economic Research.

Aguirre, J. (2021). Long-term effects of grants and loans for vocational education. *Journal of Public Economics 204*, 104539.

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

Altonji, J. G. and L. M. Segal (1996). Small sample bias in gmm estimation of covariance structures. *Journal of Business and Economic Statistics 14*(3), 353–366.

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

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review 110*(1), 48–85.

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

Banks, J., G. D. Batty, J. Breedvelt, K. Coughlin, R. Crawford, M. Marmot, J. Nazroo, Z. Oldfield, N. Steel, A. Steptoe, M. Wood, and P. Zaninotto (2021). English longitudinal study of ageing: Waves 0-9, 1998-2019.

Becker, G. S., S. D. Kominers, K. M. Murphy, and J. L. Spenkuch (2018). A theory of intergenerational mobility. *Journal of Political Economy 126*(S1), S7–S25.

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

Blanden, J., M. Doepke, and J. Stuhler (2022, April). Educational inequality. Working Paper 29979, National Bureau of Economic Research.

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

Blandin, A. and C. Herrington (2022). Family heterogeneity, human capital investment, and college attainment. *American Economic Journal: Macroeconomics 14*(4), 438–78.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital, and welfare reform. *Econometrica 84*(5), 1705–1753.

Bolt, U., E. French, J. Hentall Maccuish, and C. O'Dea (2021). The intergenerational elasticity of earnings: Exploring the mechanisms. Discussion Paper DP15975, CEPR.

Browning, M., L. P. Hansen, and J. J. Heckman (1999). Micro data and general equilibrium models. *Handbook of macroeconomics 1*, 543–633.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics 21*(3), 339–353.

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

Caucutt, E. M., L. Lochner, J. Mullins, and Y. Park (2020, September). Child skill production: Accounting for parental and market-based time and goods investments. Working Paper 27838, National Bureau of Economic Research.

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

Charles, K. K. and E. Hurst (2003). The correlation of wealth across generations. *Journal of Political Econ-*

omy 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.

CNN (2022). Student loan forgiveness polling. Accessed: 2025-06-26.

Colas, M., S. Findeisen, and D. Sachs (2021). Optimal need-based financial aid. *Journal of Political Economy 129*(2), 492–533.

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, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education 1*, 697–812.

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

Daruich, D. (2022). The macroeconomic consequences of early childhood development policies. Mimeo.

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.

Delaney, J. (2019). Born unequal? understanding the components of life time inequality. Technical report, ESRI working paper.

Duffie, D. and K. J. Singleton (1993). Simulated moments estimation of markov models of asset prices. *Econometrica 61*(4), 929–952.

French, E. (2005, 04). The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies 72*(2), 395–427.

Fu, C., S. Ishimaru, and J. Kennan (2024). Government expenditure on the public education system. *International Economic Review 65*(1), 43–73.

Fuchs-Schündeln, N., D. Krueger, A. Ludwig, and I. Popova (2022). The long-term distributional and welfare effects of covid-19 school closures. *The Economic Journal 132*(645), 1647–1683.

Gayle, G.-L., L. Golan, and M. A. Soytas (2022). What is the source of the intergenerational correlation in earnings? *Journal of Monetary Economics 129*, 24–45.

Gu, S. and L. Zhang (2024). The macroeconomic consequences of competition for college admissions. Technical report.

Heckman, J. J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of economic dynamics 1*(1), 1–58.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the highscope perry preschool program. *Journal of public Economics 94*(1-2), 114–128.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annual Review of Economics 6*(1), 689–733.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina (2008). The inheritance of educational inequality: International comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy 7*(2).

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

Jordà, Ò., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor (2019). The rate of return on everything, 1870–2015. *The Quarterly Journal of Economics 134*(3), 1225–1298.

Krueger, D., A. Ludwig, and I. Popova (2025). Shaping inequality and intergenerational persistence of poverty: Free college or better schools? *Journal of Monetary Economics 150*, 103694.

Lee, S. Y. and A. Seshadri (2019). On the intergenerational transmission of economic status. *Journal of Political Economy 127*(2), 855–921.

Lochner, L. and A. Monge-Naranjo (2016). Student loans and repayment: Theory, evidence, and policy. In

Handbook of the Economics of Education, Volume 5, pp. 397–478. Elsevier.

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.

Mullins, J. (2022). Designing cash transfers in the presence of children's human capital formation. Working Paper 2022-019, HCEO.

Newey, W. K. and D. L. McFadden (1994). Large sample estimation and hypothesis testing. In R. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics, Volume 4*. Elsevier, Amsterdam.

Nicoletti, C. and V. Tonei (2020). Do parental time investments react to changes in child's skills and health? *European Economic Review 127*, 103491.

Pakes, A. and D. Pollard (1989). Simulation and the aysmptotics of optimization estimators. *Econometrica 57*(5), 1027–1057.

Powell, J. (1994). Estimation of semiparametric models. In R. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics, Volume 4*. Elsevier, Amsterdam.

Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge 26*.

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

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters 20*(2), 177–181.

Waters, T. (2017). Taxben: The IFS tax and benefit microsimulation model. *Institute for Fiscal Studies User Guide, https://www.ifs.org.uk/uploads/publications/docs/taxben_guide.pdf*.

Yannelis, C. and G. Tracey (2022). Student loans and borrower outcomes. *Annual Review of Financial Economics 14*(1), 167–186.

Yum, M. (2022). Parental time investment and intergenerational mobility. *International Economic Review 127*(2), 855–921.

Appendixes For Online Appendix Only

A Parameter definitions

Table 13 summarizes the parameters that enter the model and which are introduced in the body of the paper (excluding the appendices).

Table 13: Parameter definitions

	Preference Parameters		**State variables**
β	Discount factor, annual	$g \in \{m, f\}$	Gender
β_{t+1}	Discount factor, between model periods	t	Model period
ν_g	Consumption weight in utility function	ed	Educational Attainment
λ	Intergenerational altruism parameter	a_t	Wealth
$1-\theta$	Share of investment time perceived as leisure	$w_{g,t}$	Wage
$\kappa_{0,t}, \kappa_{1,t}$	Time to investment conversion parameters	$h'_{t'}$	Child's skill at t'
	Labor market		**Household choices**
y_t	Household income	$c_{g,t}$	Consumption
$\tau(.)$	Net-of-tax income function	$l_{g,t}$	Leisure
$e_{g,t}$	Earnings	$hrs_{g,t}$	Work hours
η_t	Wage innovation	$ti_{g,t}$	Time investment in children
σ_η^2	Variance of wage innovation	x_t	Cash transfer ($t = 10$)
δ_j	Wage profile parameters		
	Human Capital		**Utility function and arguments**
$h'_{t'}$	Child's skill at t'	$u()$	Single period utility function
γ_j	Skill production parameters	$V_t(\mathbf{X_t})$	Value function
u_h	Stochastic skill component	$\mathbf{X_t}$	Vector of all state variables
		n_t	Number of equiv. adults in household
	Assets	T	Time endowment
$(1+r_t)$	Gross interest rate, between model periods	$\mathbf{d_t}$	Vector of decision variables
r	Annual interest rate		
			Other
	Measurement Systems	τ	Length (years) of period t
ω	Vector of child skill and time investment	$Q_g()$	Marriage probability function
		s_{t+1}	Survival rate across period t

B Time Periods, States, Choices and Uncertainty

Table 14 lists all model time periods, parents' and chilrens' age in those time periods, the state variables, choice variables, and sources of uncertainty during those time periods.

Table 14: Model time periods, and states, choices and sources of uncertainty during those time periods

Time Periods																				
Model period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Parent generation's age	0	7	11	16	23	26	33	37	42	49	55	60	65	70	75	80	85	90	95	100
Child generation's age						0	7	11	16	23										
Parent generation's datasets																				
NCDS					x		x		x	x	x									
Time use survey						x	x	x												
ELSA										x										
Child generation's datasets																				
NCDS						x	x	x												
Parent generation's states																				
Assets					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Wage of male and female					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Education of male and female					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Children's gender						x	x	x	x	x										
Children's skill						x	x	x	x	x										
Children's education										x										
Parent generation's choices																				
Work hours of male and female					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Time spent with children, male and female						x	x	x												
Consumption, male and female					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Cash transfer to children										x										
Education of children									x											
Parent generation's uncertainty																				
Wage shock of male and female					x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Initial skill of children						x														
Skill shock to children							x	x	x											
Children's partner										x										
Children's initial wage										x										
Mortality										x	x	x	x	x	x	x	x	x	x	x

Notes: Between periods 1 and 4 the parent generation makes no choices, and in this sense has no state variables or uncertainty.

C Data

We use data from the National Child Development Survey, the English Longitudinal Study of Ageing, and the UK Time Use Survey in our analysis, and use sample selection rules which are consistent across the three data sets. The sample selection rules are described in more detail below.

C.1 National Child Development Survey

Our main data set is the National Child Development Survey (NCDS), produced by the Centre for Longitudinal Studies (2017), which started with 18,558 individuals born in one week in March 1958. We use the NCDS Data in three different ways: First, for estimating the skill production functions. Second, for estimating the income process. And third, to derive moment conditions on marital matching, education shares, employment rates, the fraction of full-time work. We explain the samples used for these three purposes in more detail below.

Production function estimation: For the production function estimation, we require individuals to have a full set of observations on all skill measures, investment measures between the ages of 0-16, parental education, and parental income (see Table 1 for a full list of measures). This reduces the original sample of 48,644 observations to 11,596 observations across the four waves considered.

Income process: For the estimation of the income process, we consider the waves collected at ages 23, 33, 42, 50, and 55, leaving out age 46 due to the survey at that age being a more limited phone-only interview. This leads to a total of 54,352 observations in adulthood. Of these, we drop all self-employed people (5,932 excluded), those who are unmarried after age 23 (7,602 excluded), those for who we only have one wage observation (9,909 excluded) leaving us with 30,909 observations. We trim wages at the top and bottom 1% for each sex and education group.

Moments: For the moments, we exclude all self-employed people (5,932 excluded), and those who are unmarried after age 23 (7,602 excluded), leaving us with a total number of observations of 40,818.

C.2 English Longitudinal Study of Ageing

We use the ELSA data, produced by Banks et al. (2021), both for asset data at age 60 which we use in our moment conditions and also for the gift and inheritance data which we use in our moment condition. ELSA is a biennial survey of those 50 and older, starting in 2002. We use data up through 2016.

Although NCDS sample members are asked about assets at age 50, it is not possible to obtain a comprehensive measure of wealth as data on housing wealth is not collected; thus we use ELSA instead. For our wealth measure, we use the sum of housing wealth including second homes, savings, investments including stocks and bonds, trusts, business wealth, and physical wealth such as land, and subtract financial debt and mortgage debt.

For the asset moment condition at age 60 we begin with 2,746 respondents who are age 60 at the time of the survey. We drop members of cohorts not born before 1950 (which excludes 1,604 observations), unmarried people (which excludes 239 observations), and the self-employed (which excludes 132 observations). Finally, we have 23 individuals who live in the same household as another ELSA member. In order to not double count these households, we exclude one observation from these multi-respondent households, resulting in 748 individuals remaining.

ELSA has high quality data on gifts and inheritances in wave 6 (collected in 2012-2013). In this wave respondents were asked to recall receipt of inheritances and substantial gifts (defined as those worth over £1,000 at 2013 prices) over their entire lifetimes. Respondents are asked their age at receipt and value for three largest gifts and three largest inheritances.[15] From our original sample of 10,601 in 2012, we drop members of cohorts not born between 1950-1959 (which excludes 7,223 observations), singles (921 excluded), and self-employed (328 excluded), resulting in 2,129 individuals remaining. Of these, we only keep individuals for whom both parents had died by the time of the survey (1107 individuals) and for who we have information on the father's education resulting in a final sample of 984 individuals.

Table 15 compares education shares and median net weekly earnings in both NCDS and ELSA. The ELSA sample has modestly higher education and lower earnings, but overall the samples match quite well.

Table 15: Sample comparison: NCDS and ELSA

	Education shares			
	Male		Female	
	NCDS	ELSA	NCDS	ELSA
Low	16%	20%	22%	26%
Medium	49%	38%	49%	40%
High	35%	43%	29%	34%
	Median net weekly earnings in £			
	Male		Female	
	NCDS	ELSA	NCDS	ELSA
Low	399	315	223	171
Medium	479	383	266	221
High	665	519	399	358

Notes: In NCDS, low education includes no educational qualification or CSE2-5, Medium education includes O-level or A-level, High education includes higher qualifications or a degree. In ELSA, low education includes no educational qualification or CSE, Medium education includes O-level or A-level, High education includes higher qualifications below a degree or a degree. Earnings are median net weekly earnings in £2013.

[15] Only 3.6% of all individuals have three or more large inheritances or bequests (Crawford 2014), so the restriction is unlikely to significantly affect our results.

C.3 UK Time Use Survey

Using the measures of parental investments in the NCDS we can construct a latent time investment index. However, the NCDS does not directly measure hours of investment time. For measuring hours of investment time we use UKTUS data from 1983 and 1987. Respondents use a time diary to record activities of their day in 144 10-minute time slots for one weekday and one weekend day. In each of these slots the respondent records their main ("main activity for each ten minute slot") and secondary activities ("most important activity you were doing at the same time"), as well as who it was carried out with. We have diaries for both parents and the children, but use only the parent diaries.

We construct our measure of time spent with children by summing up across both parents the ten minute time slots during which an investment activity with a child takes place either as a main or a secondary activity. Although we know the number of children and the age of each child within the household, we do not know the precise age of the child that received the investment; we assume this to be the youngest child. We include all of the following activities as time spent with the child when constructing the investment measure: teaching the child, reading/playing/talking with child, travel escorting to/from education.

Our original sample includes 21,401 diary entries. We keep only households with two children (which excludes 15,539 observations), exclude adult children (934 observations), keep only married individuals (which excludes 531 observations), and drop entries with missing diary information (484 observations) This leaves us with 3,913 remaining observations.

D Estimation of the Skill Production Function, Parental Investment Function, and Wage Function

D.1 Production Function

The production function for skills that we estimate is as specified in equation (12) in the main text:

$$\ln h'_{t'+1} = \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'} \ln inv_{t'} + \alpha_{3,t'} \ln inv_{t'} \cdot \ln h'_{t'} + \alpha_{4,t'} ed^m + \alpha_{5,t'} ed^f + u'_{h,t'} \tag{18}$$

where $u'_{h,t'}$ is independent of all other right hand side variables.

D.2 Measurement

We do not observe children's skills (h'), or parental investments (inv) directly. However we observe $j = \{1, ..., J_{\omega,t}\}$ error-ridden measurements of each. These measurements have arbitrary scale and location.

That is for each $\omega \in \{\ln h', \ln inv\}$ we observe:

$$Z_{\omega,t,j} = \mu_{\omega,t,j} + \lambda_{\omega,t,j} \ln \omega_t + \epsilon_{\omega,t,j} \tag{19}$$

All other variables are assumed to be measured without error.

D.3 Assumptions on Measurement Errors and Shocks

Measurement errors are assumed to be independent across measures and across time. Measurement errors are also assumed to be independent of the latent variables (skill and investment), the structural shocks, and parental education $(u'_{h,t'}, ed_f, ed_m)$.

D.4 Normalizations

As mentioned above, skills and investments do not have a fixed location or scale which is why we need to normalize them. In the first period, we normalize the mean of the latent factors to be zero which fixes the location of the latent factors. In all other periods, the mean of the latent factor for skills h_t is allowed to be different from zero although the mean of investment is assumed 0 in all periods. Moreover, for each period, we set the scale parameter $\lambda_{\omega,t,1} = 1$ for one normalizing measure $Z_{\omega,t,1}$.

Agostinelli and Wiswall (2016) have shown that renormalization of the scale parameter $\lambda_{\omega,t,1} = 1$ can lead to biases in the estimation of coefficients in the case of overidentification of the production function coefficients when assuming that $\alpha_{1,t'} + \alpha_{2,t'} + \alpha_{3,t'} = 1$ in equation (18). This is not the case in our estimation as we do not assume $\alpha_{1,t'} + \alpha_{2,t'} + \alpha_{3,t'} = 1$ when estimating equation (18).

D.5 Initial Conditions Assumptions

Children are born in period $t' = 1$. The mean of h'_1, ed_f, ed_m and inv_1 are 0 by normalization and without loss of generality. h'_1 depends on parents' education and is normally distributed conditional on parents' education.

D.6 Estimation

1. **Scale parameters (λs) and variance of latent factors**. Here we estimate the parameters for the measurement equations for the child skill and investment latent variables. Using equation (19) we can derive the variance of each of the latent factors:

$$Cov(Z_{\omega,t,j}, Z_{\omega,t,j^*}) = \lambda_{\omega,t,j}\lambda_{\omega,t,j^*}\, Var(\ln \omega_t) \tag{20}$$

We have normalised $\lambda_{h,t,1} = \lambda_{inv,t,1} = 1$ to set the scale of $\ln h_t$ and of $\ln inv_t$. For each other measure $j \neq 1$, and for $\omega \in \{h', inv\}$, using equation (20) it can be shown that:

$$\lambda_{\omega,t,j} = \frac{Cov(Z_{\omega,t,j}, Z_{\omega,t,j^*})}{Cov(Z_{\omega,t,1}, Z_{\omega,t,j^*})} \tag{21}$$

The model defined in equation (21) is overidentified if we have more than three measures since there are many different combinations of j and j^* that can be used here ($j^* \neq j$). We use GMM with an identity weighting matrix to estimate the λs where the moments are all the combinations of measures possible using equation (21). With these estimates of the λs in hand, we then estimate $Var(\ln \omega_t)$ using equation (20). This equation is also overidentified with more than three measures, and again we estimate this using GMM.[16]

2. **Location parameters (μs) in measurement equations** At the child's birth ($t' = 1$), we normalize the mean of log h'_1 and inv_1 to zero. Therefore:

$$\mu_{\ln h',1,j} = \mathbb{E}[Z_{\ln h',1,j}], \quad \mu_{\ln inv,1,j} = \mathbb{E}[Z_{\ln inv,1,j}] \tag{22}$$

3. **Calculation for next step** For each measure, we need to calculate a residualized measure of each Z for $\omega_t \in \{\ln h_t, \ln inv\}$:

$$\tilde{Z}_{\omega,t,j} = \frac{Z_{\omega,t,j} - \mu_{\omega,t,j}}{\lambda_{\omega,t,j}} \tag{23}$$

This will be used below in Step 4. Note that:

$$\ln \omega_t = \tilde{Z}_{\omega,t,j} - \underbrace{\frac{\epsilon_{\omega,t,j}}{\lambda_{\omega,t,j}}}_{\equiv \tilde{\epsilon}_{\omega,t,j}} \tag{24}$$

It shows that each latent factor (skill and investment) are equal to a measure of the relevant factor plus an error, rescaled to match the scale of the factor.

4. **Estimate latent skill production technology**

[16] Note that at age 0 (period t'= 1) and age 16 (period t'= 4), we only have 2 measures of skill, respectively. Dropping logs for conciseness, to identify $\lambda_{h',4,j}$, we use covariances across time. For example, we use $Cov(Z_{h',3,j}, Z_{h',4,j}) = \lambda_{h',3,j}\lambda_{h',4,j}Cov(h'_3, h'_4)$ and $Cov(Z_{h',3,j}, Z_{h',4,j*}) = \lambda_{h',3,j}\lambda_{h',4,j*}Cov(h'_3, h'_4)$, thus $\frac{Cov(Z_{h',3,j}, Z_{h',4,j})}{Cov(Z_{h',3,j}, Z_{h',4,j*})} = \frac{\lambda_{h',3,j}\lambda_{h',4,j}Cov(h'_3, h'_4)}{\lambda_{h',3,j}\lambda_{h',4,j*}Cov(h'_3, h'_4)}$. For the normalizing measure $Z_{h',4,j*}$, $\lambda_{h',4,j*} = 1$, so this becomes $\frac{Cov(Z_{h',3,j}, Z_{h',4,j})}{Cov(Z_{h',3,j}, Z_{h',4,j*})} = \lambda_{h',4,j}$.

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

$$\ln h'_{t'+1} = \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'} \ln inv_{t'} + \alpha_{3,t'} \ln inv_{t'} \cdot \ln h_{t'} + \alpha_{4,t'} ed_m + \alpha_{5,t'} ed_f + u'_{h,t'}$$

and using equation (24) note that we can rewrite the above equation as (where we suppress logs for conciseness):

$$\begin{aligned} \frac{Z_{h',t'+1,j} - \mu_{h',t'+1,j} - \epsilon_{h',t'+1,j}}{\lambda_{h',t'+1,j}} &= \alpha_{1,t'}(\tilde{Z}_{h',t',j} - \tilde{\epsilon}_{h',t',j}) + \\ & \alpha_{2,t'}(\tilde{Z}_{inv,t',j} - \tilde{\epsilon}_{inv,t',j}) + \\ & \alpha_{3,t'}(\tilde{Z}_{inv,t',j} - \tilde{\epsilon}_{inv,t',j}) \cdot (\tilde{Z}_{h',t',j} - \tilde{\epsilon}_{h',t',j}) + \\ & \alpha_{4,t'} ed_m + \alpha_{5,t'} ed_f + \\ & u'_{h,t'} \end{aligned} \tag{25}$$

or

$$\begin{aligned} \frac{Z_{h',t'+1,j} - \mu_{h',t'+1,j}}{\lambda_{h',t'+1,j}} &= \alpha_{1,t'} \tilde{Z}_{h',t',j} + \\ & \alpha_{2,t'} \tilde{Z}_{inv,t',j} + \\ & \alpha_{3,t'} \tilde{Z}_{inv,t',j} \cdot \tilde{Z}_{h',t',j} + \\ & \alpha_{4,t'} ed_m + \alpha_{5,t'} ed_f + \\ & \Bigg(u'_{h,t'} - \tilde{\epsilon}_{h',t',j} - \tilde{\epsilon}_{inv,t',j} - \tilde{\epsilon}_{inv,t',j} \cdot \tilde{\epsilon}_{h',t',j} + \frac{\epsilon_{h',t'+1,j}}{\lambda_{h',t'+1,j}} \\ & -\alpha_{3,t'}(\tilde{Z}_{inv,t',j}\tilde{\epsilon}_{h',t',j} + \tilde{Z}_{h',t',j}\tilde{\epsilon}_{inv,t',j}) \Bigg). \end{aligned} \tag{26}$$

OLS is inconsistent here, as $\tilde{Z}_{h',t',j}$ and $\tilde{\epsilon}_{h',t',j}$ are correlated. We resolve this issue by instrumenting for $\tilde{Z}_{h',t',j}$ using the other measures of skill $\tilde{Z}_{h',t',j^*}$ in that period.

Recall that we only normalized the location of factors in the first period, but have not done so for the subsequent periods (in this case $\mu_{h',t'+1,j}$). We estimate the location parameter for each measure j by estimating a version of equation (26) with a dependent variable of just the measure $Z_{h',t'+1,j}$. The intercept then identifies $\mu_{h',t'+1,j}$.

We estimate all location parameters (the μs) and the α parameters jointly using equation (26) by using a system GMM with diagonal weights. By using the system GMM we make efficient use of all available measures.

5. **Estimate the variance of the production function shocks**

   The variance of the structural skills shock can be obtained using residuals from equation (26), by defining: $\pi_{h',t,j} \equiv \Big(u'_{h,t'} - \tilde{\epsilon}_{h',t',j} - \tilde{\epsilon}_{inv,t',j} - \tilde{\epsilon}_{inv,t',j} \cdot \tilde{\epsilon}_{h',t',j} + \frac{\epsilon_{h',t'+1,j}}{\lambda_{h',t'+1,j}} - \alpha_{3,t'}(\tilde{Z}_{inv,t',j}\tilde{\epsilon}_{h',t',j} + \tilde{Z}_{h',t',j}\tilde{\epsilon}_{inv,t,j}) \Big)$

   and using the fact that:

$$Cov\left(\frac{\pi_{h',t',j}}{\lambda_{h',t',j}}, \tilde{Z}_{h',t',j^*}\right) = \sigma^2_{h',t',j} \text{ for } j \neq j^*$$

   which is true since the measurement errors $\tilde{\epsilon}$ are uncorrelated across measures and so $Cov(\tilde{\epsilon}_{h',t',j}, \tilde{Z}_{h',t',j^*}) = 0$ if $j \neq j^*$. As before, these covariances are overidentified, so we estimate these variances using GMM where the variance covariance matrix of the $\widehat{\sigma^2_{h',t',j}}$s is estimated using the bootstrap.

E Initial Skill

Initial skill at birth is a function of mother's education level, father's education level, and a shock. Using minimum distance methods, we estimate initial skill (after adjusting for their different scales) as a function of parental education dummies. We then estimate the variance of the shock analogously to Step 5 in the previous section. Table 16 shows the results of the minimum distance, and the variance of the initial skill shock.

Table 16: Initial skill regression

	Coefficient	SE
Mother's education		
Medium	0.092	(0.041)
High	0.079	(0.103)
Father's education		
Medium	0.066	(0.044)
High	-0.007	(0.081)
Constant	-0.038	(0.022)
Variance of shock	0.880	

F Signal to Noise Ratios

Note that using equation (19) the variance of measure $Z_{\omega,t,j} = (\lambda^2_{\omega,t,j})Var(\ln \omega_t) + Var(\epsilon_{\omega,t,j})$, where $(\lambda^2_{\omega,t,j})Var(\ln \omega_t)$ comes from the variability in the signal in the measure and $Var(\epsilon_{\omega,t,j})$ represents measurement error, or "noise". The signal to noise ratios for measure $Z_{\omega,t,j}$ is calculated in the following way:

$$s_{\omega,t,j} = \frac{(\lambda^2_{\omega,t,j})Var(\ln \omega_t)}{(\lambda^2_{\omega,t,j})Var(\ln \omega_t) + Var(\epsilon_{\omega,t,j})}$$

Intuitively, this is the variance of the latent factor (signal) to the variance of the measure (signal+noise) and thus describes the information content of each measure.

Table 17 presents signal to noise ratios for skill. At birth, birthweight is the most informative measure. At age 7, reading, maths, coping, and drawing scores are all roughly equally informative. At ages 11 and 16, maths scores become the most informative.

Table 17: Signal to noise ratios: Skill measures

Age 0		*Age 7*		*Age 11*		*Age 16*	
birthweight	0.862	read	0.385	read	0.555	read	0.570
gestation	0.140	maths	0.335	maths	0.942	maths	0.713
		copy	0.259	copy	0.104		
		draw	0.281				

Note: At ages 0 and 16, we only have 2 measures of skill. Footnote 16 explains identification of the scaling parameters in this case.

Table 18 presents signal to noise ratios for investment. Here we have many measures of investment. The most informative measures when young are the frequency of father's outings with the child, and both mother's and father's frequency of reading to the child. At older ages, the most informative variable is the teacher's assessment of each parent's interest in the child's education.

Table 18: Signal to noise ratios: Investment measures

Age 0-6		*Age 7-10*		*Age 11-15*	
mum: interest	0.164	mum: interest	0.356	mum: interest	0.796
mum: outing	0.270	mum: outings	0.235	dad: interest	0.765
mum: read	0.456	dad: outings	0.166	other index	0.344
dad: outing	0.773	dad:interest	0.386	parental ambition	0.221
dad: interest	0.082	dad:role	0.033		
dad:read	0.539	parents initiative	0.206		
dad: large role	0.069	parents ambition uni	0.093		
other index	0.136	parents ambition school	0.249		
		library member	0.253		

Notes: All investment measures are retrospective, so age 0-6 investments are measured at age 7, age 7-10 investments are measured at age 11, age 11-15 investments are measured at age 16.

G Estimating the Wage Equation, Accounting for Measurement Error in Skill Levels and Wages

We estimate the wage equation laid out in equations (4) and (5), but allow for i.i.d. measurement error in wages u_t. Using those equations and noting that $v_t = \delta_5 h_4 + \sum_{k=5}^{t} \eta_k$ yields:

$$\ln w_t^* = \ln w_t + u_t = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + \delta_5 h_4 + \sum_{k=5}^{t} \eta_k + u_t \qquad (27)$$

for each gender and education group.

In our procedure we must address three issues. First, wages are measured with error u_t. Second, skill h_4 is measured with error. Third, we only observe the wage for those who work, which is a selected sample.

We address issues of selectivity by relying on our panel data as much as is possible. To address the issue of composition bias (the issue of whether lifetime high or low wage individuals drop out of the labor market first), we use a fixed effects estimator. Given our assumption of a unit root in $v_t = \delta_5 h_4 + \sum_{k=5}^{t} \eta_k$, which we estimate to be close to the truth (see Appendix G.1 for estimates that relax this assumption and allow v_t follow an AR(1)), we can allow v_5 (the first shock to wages, age 23) to be correlated with other observables, and estimate the model using fixed effects. In particular, the procedure is:

Step 0: From equation (27) note that:

$$\ln w_t^* = \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + FE + \xi_t$$

where FE is a person specific fixed effect capturing the time invariant factors $\delta_5 h_4 + \eta_5$ and ξ_t is a residual.

Step 1: Estimate δ_1, δ_2, δ_3, δ_4 using fixed effects (FE) regression.

Step 2: Predict the fixed effect:

$$\begin{aligned} \widehat{FE} &\equiv \bar{\ln w_t^*} - \hat{\delta}_1 \bar{t} - \hat{\delta}_2 \bar{t^2} - \hat{\delta}_3 \bar{t^3} - \hat{\delta}_4 \bar{PT}_t \\ &= \delta_0 + \delta_5 h_4 + \eta_5 \\ &= \delta_0 + \delta_5 \tilde{Z}_{h,4,j} + \eta_5 - \delta_5 \tilde{\epsilon}_{h,4,j} \qquad (28) \end{aligned}$$

where the means are over all observations of an individual, e.g., $\bar{t}$ is the mean age of an individual over all years she was observed, and $h_4 = \tilde{Z}_{h,4,j} - \tilde{\epsilon}_{h,4,j}$ and where $\tilde{Z}_{h,4,j}$ and $\tilde{\epsilon}_{h,4,j}$ have been defined in equation (24). The above equation holds for all measures j. Although the estimated fixed effect, $\widehat{FE}$, is affected by variability in the sequence of wage shocks $\{\eta_t\}_{t=5}^{12}$ and measurement errors $\{u_t\}_{t=5}^{12}$, this merely adds in measurement error on the left hand side variable in equation (28). However, measurement error on the right hand side variable (h_4) is more serious: we only have the noisy proxies $\tilde{Z}_{h,4,j}$ which are correlated with $\tilde{\epsilon}_{h,4,j}$ by construction. We address this problem in the next step.

Step 3: Using GMM, we project the predicted fixed effect ($\widehat{FE}$) on each measure of skill, $\tilde{Z}_{h,4,j}$, and instrument by using the respective other measures, $\tilde{Z}_{h,4,j'}$, to obtain $\hat{\delta_0}$ and $\hat{\delta_5}$. Since we have two measures of skill (reading and math), we have two equations and two instruments. When reading is the skill measure, we instrument for this using math, and vice versa. Our GMM procedure efficiently combines different measures of skill and yields consistent estimates of $\hat{\delta_0}$ and $\hat{\delta_5}$ even in the presence of measurement error in the skill measures.

Step 4: Then use covariances and variances of residuals to calculate shock variances.

Substituting a noisy measure of skill into the wage equation (27) yields

$$\ln w_t^* = \delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + \delta_5 \tilde{Z}_{h,4,j} + \sum_{k=5}^{t} \eta_k + u_t - \delta_5 \tilde{\epsilon}_{h,4,j}$$

where we use the fact that $h_4 = \tilde{Z}_{h,4,j} - \tilde{\epsilon}_{h,4,j}$ as defined in equation (24). Next we define a wage residual that will exist for each skill measure:

$$\widetilde{\ln w_{t,j}} \equiv \ln w_t^* - (\delta_0 + \delta_1 t + \delta_2 t^2 + \delta_3 t^3 + \delta_4 PT_t + \delta_5 \tilde{Z}_{h,4,j}) = \sum_{k=5}^{t} \eta_k + u_t - \delta_5 \tilde{\epsilon}_{h,4,j}$$

Note that from the measurement equation (19), $Var(\tilde{Z}_{h,4,j}) = Var(h_4) + Var(\tilde{\epsilon}_{h,4,j})$, where we have previously estimated $Var(h_4)$ using equation (20) and $Var(\tilde{Z}_{h,4,j})$ is the variance of the renormalized measures in the data. We can then back out the variance of the measurement error and plug it into the following equation to estimate the parameters of the wage shocks:

$$Cov(\widetilde{\ln w_{t,j}}, \widetilde{\ln w_{t+l,j}}) = Var(\sum_{k=5}^{t} \eta_k) + \delta_5^2 Var(\tilde{\epsilon}_{h,4,j}) \text{ for } l > 0$$

$$Var(\widetilde{\ln w_{t,j}}) = Var(\sum_{k=5}^{t} \eta_k) + Var(u_t) + \delta_5^2 Var(\tilde{\epsilon}_{h,4,j})$$

Step 5: Correct the δ parameters for selection. The fixed-effects estimator is identified using wage growth for workers. If wage growth rates for workers and non-workers are the same, composition bias problems—the question of whether high wage individuals drop out of the labor market later than low wage individuals—are not a problem. However, if individuals leave the market because of a wage drop, such as from job loss, then wage growth rates for workers will be greater than wage growth for non-workers. This selection problem will bias estimated wage growth upward.

We control for selection bias by finding the wage profile that, when fed into our model, generates the same fixed effects profile as in the estimates using the NCDS data. Because the simulated fixed effect profiles are computed using only the wages of those simulated agents that work, the profiles should be biased upwards for the same reasons they are in the data. We find this bias-adjusted wage profile using the iterative procedure described in French (2005).

G.1 Wage shock process estimates without imposing random walk

In Section 5.2, we impose that wage shocks have an autocovariance of 1. Table 19 shows the coefficients and standard errors when we relax this assumption and allow the persistence parameter to be different from one. We also report results from an overidentification test statistic. To do this we initially regress log wages on age, education, skill and part time status as before, then estimate the process for the residuals using an error components model where we match the variance covariance matrix of wage residuals. When estimating, we allow for an AR(1) process with homoskedastic (i.e., with age-invariant variances) innovations and a transitory shock in which we allow for heteroskedasticity. We have 5 periods of data, and thus 15 unique elements of the variance covariance matrix which we treat as moment conditions for each gender/education group. We estimate the variances of the transitory shocks (5 parameters), the initial variance of the AR(1) component, the variance of the AR(1) shocks, and ρ, meaning that we have 8 parameters to estimate and thus 15-8=7 degrees of freedom, meaning that under the null of correct model specification our test statistic should be distributed $\chi^2(7)$. Overall, the model fits the data well and we cannot reject the hypothesis of correct model specification for many groups. Perhaps more importantly, we can see that for all groups except low educated females, we cannot reject the hypothesis that the persistence parameter is 1. Even for this group, the value of $\rho = 0.94$. Thus throughout we assume $\rho = 1$ for all groups.

Table 19: Estimates for AR(1) process without random walk restriction

		Male	
Education:	Low	Medium	High
ρ	1.034	0.968	1.027
	(0.022)	(0.018)	(0.019)
Test stat:	10.74	12.96	36.60
		Female	
Education:	Low	Medium	High
ρ	0.940	0.985	0.971
	(0.029)	(0.023)	(0.023)
Test stat:	35.38	23.56	19.54

This table shows the persistency parameter for an AR(1) wage shock process when we relax the assumption that the process is a random walk. Bootstrapped standard errors are in parentheses. The rows entitled "Test stat" show the overidentification test statistic.

H Computational Details

This Appendix details how we solve for optimal decision rules as well as our simulation procedure. We describe solving for optimal decision rules first.

1. To find optimal decision rules, we solve the model backwards using value function iteration. The state variables of the model are model period, assets, wage rates, education levels, childrens' gender, childrens' skill, and childrens' education. At each model period, we solve the model for 50 grid points for assets, 10 grid points for wage rates for each spouse, 3 education levels for each spouse, childrens' gender, childrens' skill (5 points), and childrens' education. Because we assume that the two children are identical, receive identical shocks, and that parents make identical decisions towards the two children, we only need to keep track of the state variables for one child. Our approach for discretizing wage shocks follows Tauchen (1986). The bounds for the discretisation of the wage process is $\pm$ 3 standard deviations. For skills we use Gauss-Hermite procedures to integrate. We use linear interpolation between grid points when on the grid, and use linear extrapolation outside of the grid.

2. Both parents choose between between 4 levels of working hours (non-employed, part-time, full-time, over-time) and in model periods $t = \{6, 7, 8\}$ they can choose between six levels of time spent with children. In all model periods except $t = 10$ we solve for the optimal level of next period assets using golden search. In period $t = 10$ parents may also transfer assets to children: we solve this two-dimensional optimization problem using Nelder-Mead. We back out household consumption from the budget constraint and then solve for individual level consumption from the intra-temporal first order condition, which delivers the share of household consumption going to the male in the household. As this first order condition is a non-linear function we approximate the solution using

the first step of a third order Householder algorithm. This allows us to use the information contained in the first three derivatives of the first order condition. We found this method to give fast and accurate solutions to the intra-temporal problem.

Next we describe our simulation procedure.

1. Our initial sample of simulated individuals is large, consisting of 50,000 random draws of individuals in the first wave of our data at age 23. Given our simulated sample of individuals is larger than the number of individuals observed in the data, most observations in the data will be drawn multiple times. We take random Monte Carlo draws of education and assets, which are the state variables that we believe are measured accurately and are observed for everyone in the data. For the variables with a large amount of measurement error, or which are not observed for all sample members (i.e., initial skill of child, and wages of each parent), we exploit the model implied joint distribution of these state variables. We assume a child's gender is randomly distributed across the population.

2. Given the optimal decision rules, the initial conditions of the state variables, and the simulated histories of shocks faced by both parents and children, we calculate life histories for savings, consumption, labor supply, time and education investments in children, which then implies histories for childrens' skill and educational attainment. For discrete choice variables (e.g. participation), we evaluate whether the choice is the same at all surrounding grid points. If so, we take the implied discrete variable, and if any of the continuous state variables (e.g. assets) is between grid-points, we interpolate to find the implied decision rule. If the implied discrete choice is not the same at all surrounding grid points, we re-solve the household's problem for the exact value of all of the state variables.

3. We aggregate the simulated data in the same way we aggregate the observed data and construct moment conditions. We describe these moments in greater detail in Appendix I. Our method of simulated moments procedure delivers the model parameters that minimize a GMM criterion function which we also describe in Appendix I.

4. To search for the parameters that minimize the GMM criterion function, we use the BOBYQUA algorithm developed by Powell (2009). This is a derivative-free algorithm that uses a trust region approach to build quadratic models of the objective function on sub-regions.

I Moment Conditions and Asymptotic Distribution of Parameter Estimates

We estimate the parameters of our model in two steps. In the first step, we estimate the vector χ, the set of parameters then can be estimated without explicitly using our model. In the second step, we use the method of simulated moments (MSM) to estimate the remaining parameters, which are contained in the $M \times 1$ parameter vector $\Delta =(\beta, \nu_f, \nu_m, \gamma, \lambda, \theta, \{\kappa_{1,t'}\}_{\{t'=1,2,3\}})$. Our estimate, $\hat{\Delta}$, of the "true" parameter vector Δ_0 is the value of Δ that minimizes the (weighted) distance between the life-cycle profiles found in the data and the simulated profiles generated by the model.

We match data from three different sources. For most of our moments we use data from the NCDS. However, the NCDS currently lacks comprehensive asset and transfer data after age 23, and does not have detailed time use information with children. Thus, for the asset and transfer data we match data from ELSA, and for the information on time with children we also use UKTUS.

From the NCDS we match, for three education groups ed, two genders (male and female) g, $T = 5$ different periods: $t \in \{5, 7, 9, 10, 11\}$ (which corresponds to ages 23, 33, 42, 50, 55) the following $3 \times 2 \times T = 6T$ moment conditions for both employment rates and mean annual work hours of workers, creating a total of $12T$ moment conditions. In addition, from the NCDS we match age 16 skill and the mean education leaving age, each conditional on father's education level (6 moment conditions).

From ELSA we match mean lifetime inter-vivos transfers received (1 moment) and also household median wealth at age 60 (1 moments).

From UKTUS we match mean annual time spent with children, by age of child (ages 0-7, 8-11, 12-16) and gender of parent ($3 \times 2 = 6$ moments).

In the end, we have a total of $J = 74$ moment conditions.

Our approach accounts explicitly for the fact that the data are unbalanced since some individuals leave the sample; furthermore, and we use multiple datasets, so an individual who belongs in one sample (e.g., NCDS) likely does not belong to another sample (e.g., ELSA or UKTUS). Suppose we have a dataset of I independent individuals that are each observed in up to J separate moment conditions. Let $\varphi(\Delta; \chi_0)$ denote the J-element vector of moment conditions described immediately above, and let $\hat{\varphi}_I(.)$ denote its sample analog. Letting $\widehat{\mathbf{W}}_I$ denote a $J \times J$ weighting matrix, the MSM estimator $\hat{\Delta}$ is given by

$$\underset{\Delta}{\operatorname{argmin}} \ \frac{I}{1+\tau} \ \hat{\varphi}_I(\Delta; \chi_0)' \widehat{\mathbf{W}}_I \hat{\varphi}_I(\Delta; \chi_0),$$

where τ is the ratio of the number of data observations to the number of simulated observations.

In practice, we estimate χ_0 as well, using the approach described in the main text. Computational concerns, however, compel us to treat χ_0 as known in the analysis that follows. Under regularity conditions stated in Pakes and Pollard (1989) and Duffie and Singleton (1993), the MSM estimator $\hat{\Delta}$ is both consistent and asymptotically normally distributed:

$$\sqrt{I}\left(\hat{\Delta} - \Delta_0\right) \rightsquigarrow N(0, \mathbf{V}),$$

with the variance-covariance matrix $\mathbf{V}$ given by

$$\mathbf{V} = (1+\tau)(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1},$$

where $\mathbf{S}$ is the variance-covariance matrix of the data;

$$\mathbf{D} = \frac{\partial \varphi(\Delta; \chi_0)}{\partial \Delta'}\Big|_{\Delta=\Delta_0} \tag{29}$$

is the $J \times M$ gradient matrix of the population moment vector; and $\mathbf{W} = \text{plim}_{I\to\infty}\{\widehat{\mathbf{W}}_I\}$.

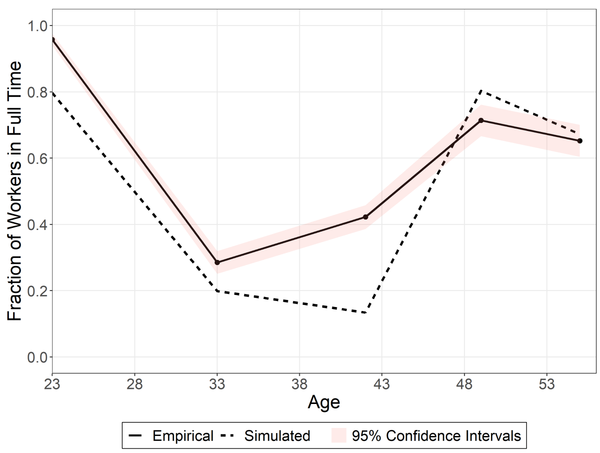
The asymptotically efficient weighting matrix arises when $\widehat{\mathbf{W}}_I$ converges to $\mathbf{S}^{-1}$, the inverse of the variance-covariance matrix of the data. When $\mathbf{W} = \mathbf{S}^{-1}$, $\mathbf{V}$ simplifies to $(1+\tau)(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}$.

But even though the optimal weighting matrix is asymptotically efficient, it can be biased in small samples. (See, for example, Altonji and Segal (1996).) We thus use a "diagonal" weighting matrix which consists of the inverse of the moments along the diagonal and 0s off this diagonal. This matrix weights more heavily moments with low means so that they too will contribute significantly to the GMM criterion function, regardless of how precisely estimated they are.

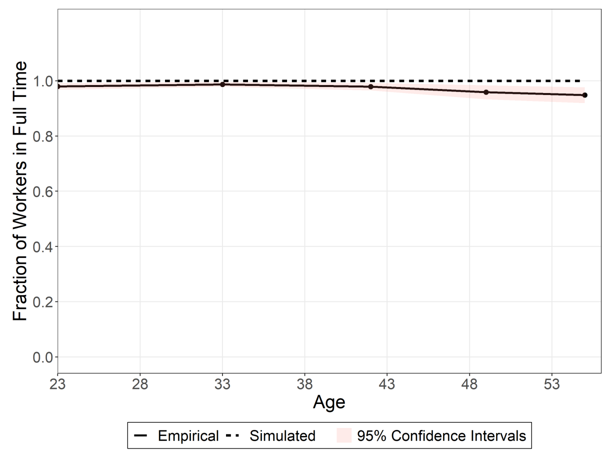
We estimate $\mathbf{D}$, $\mathbf{S}$, and $\mathbf{W}$ with their sample analogs. For example, our estimate of $\mathbf{S}$ is the $J \times J$ estimated variance-covariance matrix of the sample data. One complication in estimating the gradient matrix $\mathbf{D}$ is that the functions inside the moment condition $\varphi(\Delta; \chi)$ are non-differentiable at certain data points (e.g., for employment). This means that we cannot consistently estimate $\mathbf{D}$ as the numerical derivative of $\hat{\varphi}_I(.)$. Our asymptotic results therefore do not follow from the standard GMM approach, but rather the approach for non-smooth functions described in Pakes and Pollard (1989), Newey and McFadden (1994), and Powell (1994). When calculating gradients we vary step-sizes, then take the average gradient over the different step-sizes.

J Further Details on Model Fit

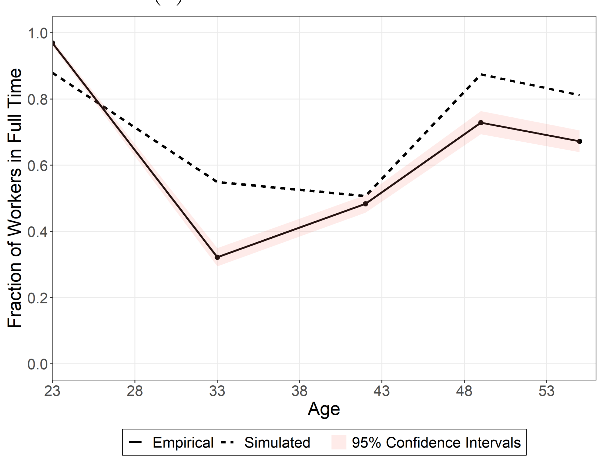
Figure 6: Model fit: full-time work conditional on employment



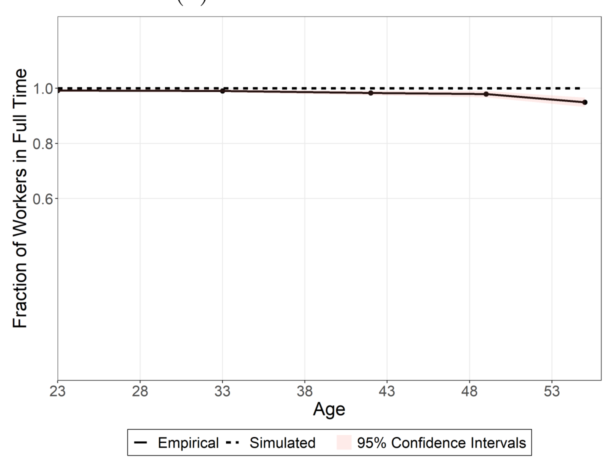
(a) Women: low educated



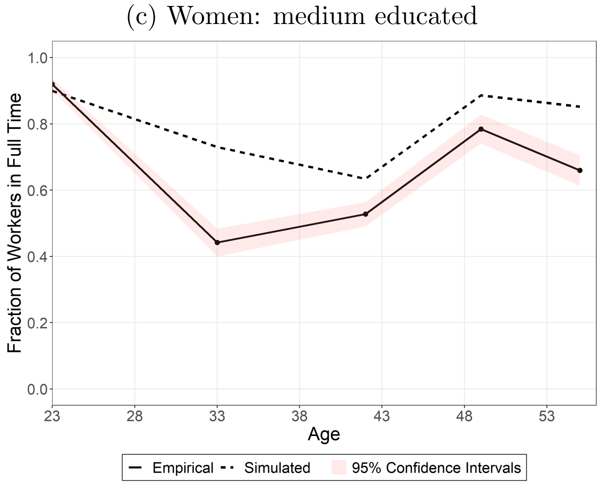
(b) Men: low educated



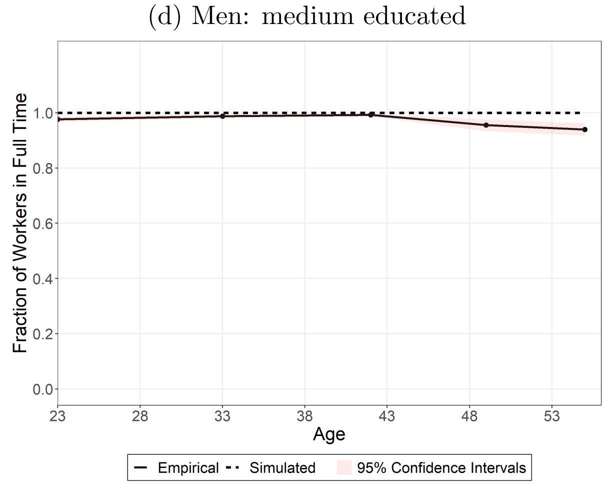
(c) Women: medium educated



(d) Men: medium educated



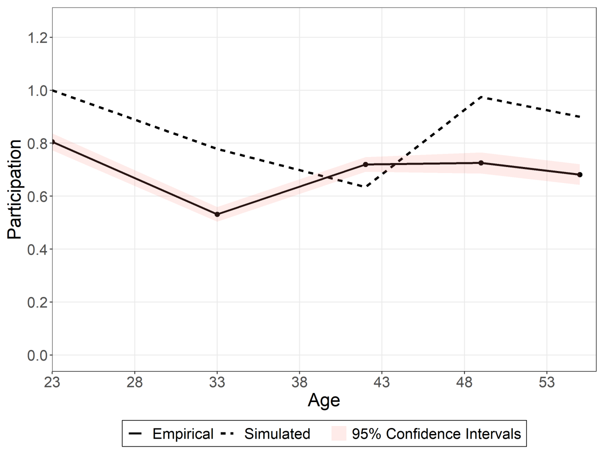
(e) Women: high educated



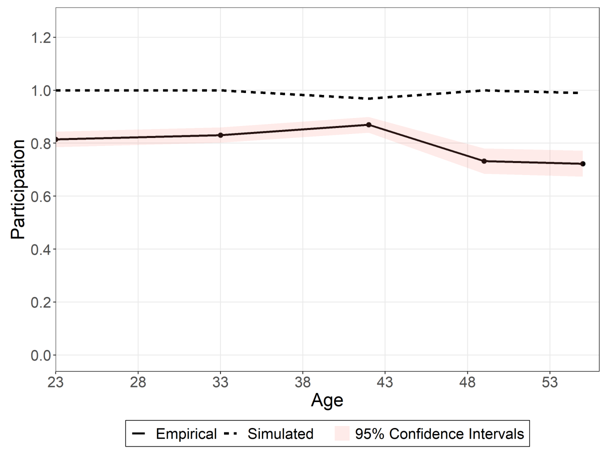
(f) Men: high educated

Notes: Figures show fraction in full time work at different ages conditional on being employed for women and men. Empirical data come from NCDS.

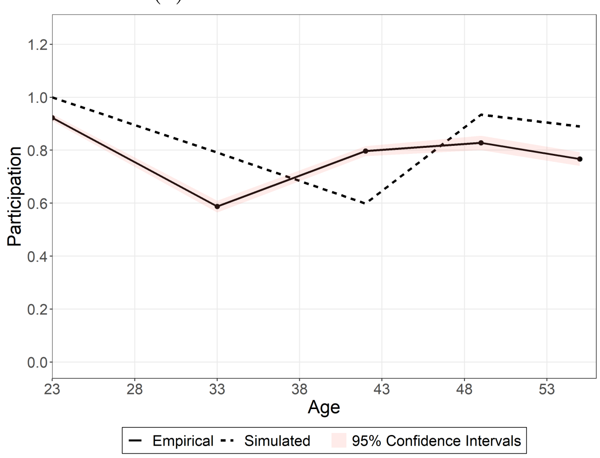
Figure 7: Model fit: participation



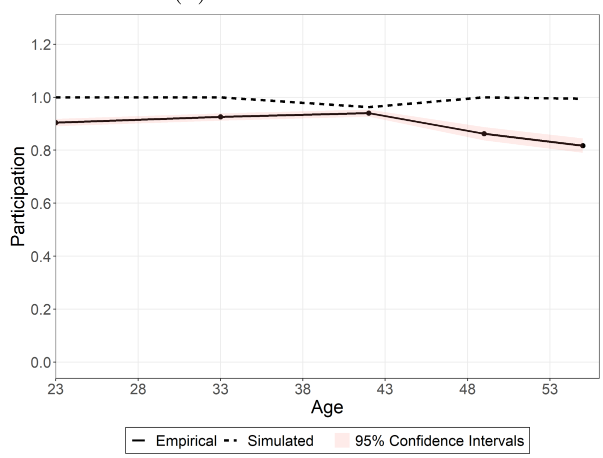
(a) Women: low educated



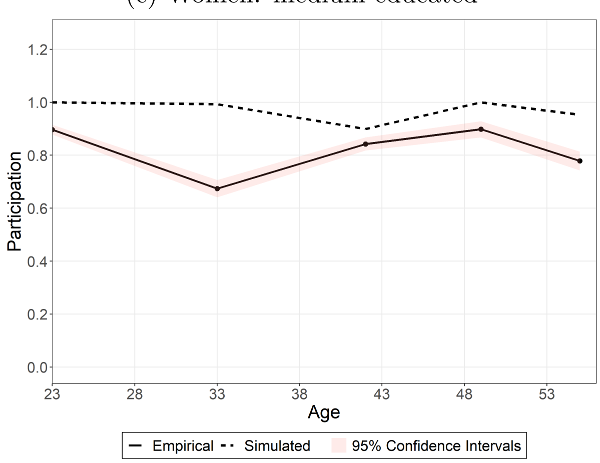
(b) Men: low educated



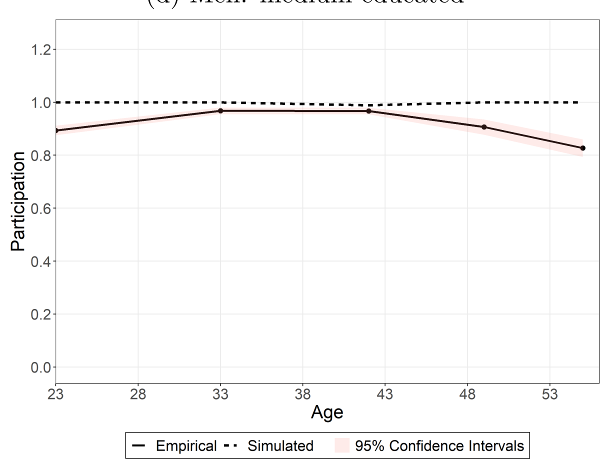
(c) Women: medium educated



(d) Men: medium educated



(e) Women: high educated



(f) Men: high educated

Notes: Figures show fraction of individuals employed at different ages for women and men. Empirical data come from NCDS.

K Identification of the time cost of investments θ

To give some intuition regarding the identification of θ, we use a simplified two period version of our dynastic model, where we abstract from couples, uncertainty, and where we assume a linear production function.

The household's state variables are: education ed, skill h, and their initial assets a_1. The parent is altruistic towards their child and incorporates their child's value function into their problem, but discounts it by factor λ. Households choose consumption c_t, leisure l_t, time investments ti_t, monetary transfers to their child x'_1 and the education of the child ed' which can be dropout (D), high school (HS) or college (C). Each education choice is associated with a price p_k, $k \in \{D, HS, C\}$, which can be interpreted as the price of foregone labor earnings of the child. The child's initial assets equal the monetary transfer from the parent.

We first describe the discrete decision problem of the parent who selects their children's education level. They maximize their value function which nests the child's value function:

$$V(ed, h, a_1) = \max_{ed'=\{D,HS,C\}} \{V_{ed'=D}, V_{ed'=HS}, V_{ed'=C}\} \tag{30}$$

where $V_{ed'=k}$ denotes the value of the problem if the parents choose education level k for their child. The above nests the following decision problem over consumption, leisure, time investments and asset transfers:

$$V_{ed'=k}(ed, h, a_1) = \max_{c_1,l_1,ti_1,x'_1,c_2,l_2,ti_2} u(c_1, l_1) + \beta u(c_2, l_2) + \lambda V'(ed' = k, h', a'_1) \tag{31}$$

subject to:

$$(1+r)^2 a_1 + (1+r) hrs_1 w_1(ab, ed) + hrs_2 w_2(h, ed) = (1+r)c_1 + c_2 + x'_1 + \sum_{ed'=\{D,HS,C\}} p_k \mathbf{1}_{[ed'=k]} \tag{32}$$

$$h' = \alpha_0 + \alpha_1 ti_1 + \alpha_2 ti_2 \tag{33}$$

$$T = \theta ti_1 + hrs_1 + l_1 \tag{34}$$

$$T = \theta ti_2 + hrs_2 + l_2 \tag{35}$$

$$a'_1 = x'_1, x'_1 \geq 0 \tag{36}$$

where (32) describes the monetary budget constraint over 2 periods, (33) shows the human capital production function over two periods where α_1, α_2 are the productivity of time investments for final skill[17]. Equations (34) and (35) are the time constraints in period 1 and 2, and (36) states that initial assets equal the initial parental cash transfer.

Assuming interior conditions for the choice variables $\{c_1, l_1, ti_1, c_2, l_2, ti_2\}$ but allowing the constraint $x'_1 \geq 0$ to bind we can now rewrite this problem and derive optimality conditions:

$$V_{ed'=k}(ed, h, a_1) = \max_{c_1,l_1,ti_1,x'_1,c_2,l_2,ti_2} u(c_1, l_1) + \beta u(c_2, l_2) + \lambda V'(h', ed', x'_1)$$
$$+\mu[(1+r)^2 a_1 + (1+r)hrs_1 w_1(h, ed) + hrs_2 w_2(h, ed) - (1+r)c_1 - c_2 - x'_1 - p_k \mathbf{1}_{[ed'=k]}]$$
$$+\kappa(\alpha_0 + \alpha_1 ti_1 + \alpha_2 ti_2 - h') + \phi(x'_1)$$
$$+\zeta_1(\theta ti_1 + hrs_1 + l_1 - T)$$
$$+\zeta_2(\theta ti_2 + hrs_2 + l_2 - T)$$

Euler equation: $\frac{\partial u}{\partial c_1} = \beta \frac{\partial u}{\partial c_2}(1+r)$

FOC wrt ti_1: $\zeta_1\theta - \kappa\alpha_1 = 0$

FOC wrt h': $\kappa + \lambda \frac{\partial V'}{\partial h'} = 0$

$$\kappa + \lambda\mu'[(1+r)\frac{\partial w'_1(h', ed'=k)}{\partial h'} hrs'_1 + \frac{\partial w'_2(h', ed'=k)}{\partial h'} hrs'_2] = 0$$

FOC wrt l_1:$-\zeta_1 + \frac{\partial u}{\partial l_1} = 0$

FOC wrt l_2:$-\zeta_2 + \beta\frac{\partial u}{\partial l_2} = 0$

FOC wrt hrs_1:$-\zeta_1 + \mu(1+r)w_1(h, ed) = 0$

FOC wrt hrs_2:$-\zeta_2 + \mu w_2(h, ed) = 0$

FOC wrt x'_1:$-\mu + \lambda\frac{\partial V'}{\partial x'_1} = 0$

$$\mu = \lambda\mu'(1+r)^2 + \phi \Rightarrow \mu \geq \lambda\mu'(1+r)^2$$

From this, we can derive the following optimality condition for investments in period 1:

$$w_1(h, ed)\theta \leq \alpha_1[\frac{1}{(1+r)^2}\frac{\partial w'_{1'}(h', ed')}{\partial h'} hrs'_{1'} + \frac{1}{(1+r)^3}\frac{\partial w_{2'}(h', ed')}{\partial h'} hrs'_{2'}] \tag{37}$$

This equation is key to understanding the identification of θ. On the left hand side, we have the marginal cost of investments to the parent which is their wage times θ. Recall that the the marginal value

[17] The κ parameters do not feature in this simplified model. They can be assumed as known, and are identified separately as shown in the next section.

of one hour of foregone leisure is equal to the wage. Furthermore, θ is the the leisure cost of an hour spent with the child. Thus the left hand side equals the value of leisure lost per hour of time spent with the child and is thus the marginal cost of parental time investments. On the right hand side, we have the marginal benefit of an hour spent with the child; this is the increase in the present discounted value of the child's future income from the hour of investment. The increase equals the productivity of an hour of time α_1, multiplied by the resulting marginal increase in income over the life cycle to the child. If cash transfers are positive equation (37) holds with equality, although if cash transfers are 0 then it is an inequality. Dividing both sides by $w_1(h, ed)$ shows that we can place an upper bound on θ by calculating the present values of the gain in child's lifetime income from one hour of time investment relative to the wage.

L Identification of κ

Our structural model maps hours of parental time into future skill. However, the NCDS has latent investments and future skill. Here we show more on the identification of κ, which maps hours of time into latent investments.

As described in section 4.1, we assume hours of parental time spent with children $ti_{m,t'}, ti_{f,t'}$ are converted to latent investment units $inv_{t'}$ according to equation (13) which we reproduce here:

$$\ln inv_{t'} = \kappa_{0,t'} + \kappa_{1,t'} \ln(\underbrace{ti_{m,t'} + ti_{f,t'}}_{ti'_t})$$

where $\kappa_{1,t'}$ is the hours-to-latent investments conversion parameter which determines the productivity of an hour of time and $\kappa_{0,t'}$ is a constant. We allow the κ parameters to vary by age. With three investment periods and two parameters in each period, this gives us six parameters to estimate.

To gain intuition regarding the identification of the κ parameters, recall equation (14) which shows the relationship between time investments and skill:

$$\begin{aligned}\ln h'_{t'+1} &= \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'} \ln inv_{t'} + \alpha_{3,t'} \ln inv_{t'} \cdot \ln h_{t'} + \alpha_{4,t'} ed^m + \alpha_{5,t'} ed^f + u'_{h,t'} \\ &= \alpha_{1,t'} \ln h'_{t'} + \alpha_{2,t'}(\kappa_{0,t'} + \kappa_{1,t'} \ln ti_{t'}) + \alpha_{3,t'}(\kappa_{0,t'} + \kappa_{1,t'} \ln ti_{t'}) \cdot \ln h'_{t'} + \alpha_{4,t'} ed^m + \alpha_{5,t'} ed^f + u'_{h,t'}\end{aligned}$$

The top line is the estimating equation using the NCDS data: the α parameters are estimated using the latent investment, skill, and parental education measures in the NCDS data. The κ parameters are estimated within the dynamic programming model. Identification of the κ_1 parameters comes from the dispersion in time investments $ti_{m,t'} + ti_{f,t'}$ (from the UKTUS data) and the corresponding dispersion in skill (from the NCDS). All the α parameters and parental education ed^m, ed^f are known. From UKTUS we know the distribution of time investments and can see a measure of dispersion around the mean, such as differences by gender, which we match. From the NCDS we know the dispersion of final skill, for example, the gradient by skill by parental education after controlling for the direct effect of parental education on skill: $\alpha_{4,t'}ed^m + \alpha_{5,t'}ed^f$.. Taking the observed dispersion in time from the UKTUS as given larger $\kappa_{1,t'}$ translates into more dispersion in final skills, and a smaller $\kappa_{1,t'}$ into less. Hence, observing both the dispersion in time inputs and resulting skills allows us to pin down the $\kappa_{1,t'}$.

M Updating the matching probabilities in counterfactuals

In Section 5.3, we show that marital matching probabilities depend on the education level of the male and the female. These probabilities reflect the prevailing distribution of education levels in the population for the cohort we study. When evaluating the education subsidy in Section 7.3, we must account for the fact that, in counterfactual settings, the distribution of education levels in the population may change, which will lead to changes in the marital matching probabilities. We account for this in our counterfactuals by allowing matching probabilities to depend on population education shares.

We estimate these matching probabilities as a function of the distribution of education levels observed in the population using data from the Family Expenditure Survey (FES) and its successor surveys from 1978 to 2017. During this time, there were major changes in the distribution of education, both for men and women. For example, the share of women with high education increased from less than 10% in 1987 to more than 40% in 2017. We use these data to estimate the following ordered logit model where for each gender and education level, we estimate the probability of matching with someone of the other gender with a certain education level, conditional on the distribution of education in the population of both genders. For example, we estimate the probability that an individual of gender g and education level $ed = j$ partners with an individual of education level $ed^P = i \in \{$low, medium, and high educated$\}$ as:

$$\begin{aligned} p_{i,j,g} = Pr\left(ed^P_{j,g} = i\right) &= Pr\left(\kappa_{i-1,j,g} < \mathbf{x}\boldsymbol{\beta}_{j,g} + u \leq \kappa_{i,j,g}\right) \\ &= \frac{1}{1+\exp\left(-\kappa_{i,j,g} + \mathbf{x}\boldsymbol{\beta}_{j,g}\right)} - \frac{1}{1+\exp\left(-\kappa_{i-1,j,g} + \mathbf{x}\boldsymbol{\beta}_{j,g}\right)} \end{aligned} \tag{38}$$

where $\mathbf{x}\boldsymbol{\beta}_j = \beta_{1,j}S_{m,low} + \beta_{2,j}S_{m,medium} + \beta_{3,j}S_{f,low} + \beta_{4,j}S_{f,med}$. $S_{g,ed}$ denotes the share in the population who are in gender group g and education group ed and the $\kappa_{i,j}$ parameters are the estimated thresholds for each group. Equation (38) is estimated separately for each education level and gender. In our dynastic model, any given policy environment generates population shares $S_{g,ed}$. These can be used with the parameters estimated here to deliver the matching probabilities that characterize the marriage market under the new equilibrium.