A Data Appendix

A.1 PSID Data and Sample Selection

We use all waves of the PSID from 1968 to 2015. The PSID started in 1968, collecting information on a sample of approximately 5,000 households. In subsequent years both the original families and their splitoffs (i.e. children who moved out of the parent household) have been followed. This is an essential feature of the data that makes it suitable for the analysis in this paper. To match parents and children we use the PSID Family Identification Mapping System, resulting in a panel of parent-child pairs. We drop pairs for which the age difference between parents and children in less than 15 years and larger than 65 years, as well as pairs with missing occupation of the child in all years.

We transform the panel of parent-child pairs into a cross-section of parent-child pairs with the following variables:

1. *Occupation*: defined, for both parents and children, as the most frequenty held occupation between age 22 and age 55. To study occupational choice and characteristics of occupations, we map detailed (and changing) occupation classifications in the PSID into 54 occupations, listed in Table 5.

Table 5: Occupation Groups, Baseline

Occ	Description	% children in occ	% parents in occ
1	Executive, Administrative, and Managerial Occupations	10.934	20.250
2	Management Related Occupations	3.494	2.372
3	Architects	0.194	0.237
4	Engineers	1.531	3.494
5	Mathematical and Computer Scientists	1.596	0.970
6	Natural Scientists	0.712	0.733
7	Health Diagnosing Occupations	0.819	1.359
8	Health Assessment and Treating Occupations	2.415	0.410
9	Therapists	0.755	0.194
10	Teachers, Postsecondary	0.863	0.884
11	Teachers, Except Postsecondary	5.823	2.308
12	Librarians, Archivists, and Curators	0.173	–
13	Social Scientists and Urban Planners	0.431	0.022
14	Social, Recreation, and Religious Workers	1.617	0.863

15	Lawyers and Judges	1.035	0.970
16	Writers, Artists, Entertainers, and Athletes	2.286	1.014
17	Health Technologists and Technicians	1.682	0.388
18	Engineering and Related Technologists and Technicians	0.690	1.035
19	Science Technicians	0.216	0.194
20	Technicians, Except Health, Engineering, and Science	1.316	1.121
21	Sales Occupations	9.144	5.844
22	Miscellaneous Administrative Support Occupations	3.235	0.323
23	Computer and Communication Equipment Operators	0.280	0.194
25	Secretaries, Stenographers, and Typists	3.429	0.194
25	Information Clerks	0.669	0.086
26	Records Processing Occupations, Except Financial	0.453	0.518
27	Financial Records Processing Occupations	1.596	0.151
28	Mail Distribution Occupations	0.496	0.712
29	Material Recording, Scheduling, and Distributing Clerks	1.531	1.423
30	Adjusters and Investigators	1.596	0.323
31	Private Household Occupations	0.819	0.151
32	Guards	0.561	0.453
33	Firefighting and Fire Prevention Occupations	0.129	0.582
34	Police and Detectives	1.251	1.596
35	Food Preparation and Service Occupations	5.435	0.474
36	Health Service Occupations	3.062	0.194
37	Cleaning and Building Service Occupations	1.617	2.027
38	Personal Service Occupations	3.451	0.474
39	Farm Operators and Managers	0.604	3.278
40	Farm and Agricultural Occupations, Except Managerial	0.474	1.078
41	Forestry, Logging, Fishing and Hunting Occupations	0.259	0.561
42	Vechicle Mechanics	2.653	4.960
43	Electronic Repairers	0.863	1.466
44	Miscellaneous Repair Occupations	0.496	1.014
45	Construction Trade Occupations	4.011	6.491
46	Extractive Occupations	0.151	0.302
47	Precision Production Supervisors	0.712	2.372
48	Precision Production Workers	1.423	3.429
49	Machine Operators	3.105	7.160

50	Fabricators	1.380	1.639
51	Production Inspectors	0.410	0.518
52	Motor Vehicle Operators	3.105	6.448
53	Non Motor Vehicle Operators	1.790	2.674
54	Freight, Stock and Material Handlers	1.229	2.070

In robustness exercises, we also consider a finer occupation classification, with the 80 occupation groups listed in Table 6.

Table 6: Occupation Groups, Robustness

Occ	Description
1	Executive, Administrative, and Managerial Occupations
2	Management Related Occupations
3	Architects
4	Engineers
5	Mathematical and Computer Scientists
6	Natural Scientists
7	Health Diagnosing Occupations
8	Health Assessment and Treating Occupations
9	Therapists
10	Teachers, Postsecondary
11	Teachers, Except Postsecondary
12	Librarians, Archivists, and Curators
13	Social Scientists and Urban Planners
14	Social, Recreation, and Religious Workers
15	Lawyers and Judges
16	Writers, Artists, Entertainers, and Athletes
17	Health Technologists and Technicians
18	Engineering and Related Technologists and Technicians
19	Science Technicians
20	Technicians, Except Health, Engineering, and Science
21	Supervisors and Proprietors, Sales Occupations
22	Sales Representatives, Finance and Business Services
23	Sales Representatives, Commodities Except Retail
24	Sales Workers, Retail, Personal Services and Sales Related Occupations

25	Supervisors, Administrative Support Occupations
26	Computer Equipment Operators
27	Secretaries, Stenographers, and Typists
28	Information Clerks
29	Records Processing Occupations, Except Financial
30	Financial Records Processing Occupations
31	Duplicating, Mail, and Other Office Machine Operators
32	Communications Equipment Operators
33	Mail and Message Distributing Occupations
34	Material Recording, Scheduling, and Distributing Clerks
35	Adjusters and Investigators
36	Miscellaneous Administrative Support Occupations
37	Private Household Occupations
38	Supervisors, Protective Service Occupations
39	Firefighting and Fire Prevention Occupations
40	Police and Detectives
41	Guards
42	Food Preparation and Service Occupations
43	Health Service Occupations
44	Cleaning and Building Service Occupations, Except Household
45	Personal Appearance Occupations
46	Recreation and Hospitality Occupations
47	Child Care Workers
48	Misc. Personal Care and Service Occupations
49	Farm Operators and Managers
50	Farm and Agricultural Occupations, Except Managerial
51	Forestry and Logging Occupations
52	Fishers, Hunters, and Trappers
53	Supervisors, mechanics and repairers
54	Vehicle and Mobile Equipment Mechanics and Repairers
55	Electrical and Electronic Equipment Repairers
56	Miscellaneous Mechanics and Repairers
57	Supervisors, Construction Occupations
58	Construction Trades, Except Supervisors
59	Extractive Occupations

60	Supervisors, Production Occupations
61	Precision Metal Working Occupations
62	Precision Woodworking Occupations
63	Precision Textile, Apparel, and Furnishings Machine Workers
64	Precision Workers, Assorted Materials
65	Precision Food Production Occupations
66	Plant and System Operators
67	Metalworking and Plastic Working Machine Operators
68	Metal and Plastic Processing Machine Operators
69	Woodworking Machine Operators
70	Printing Machine Operators
71	Textile, Apparel, and Furnishings Machine Operators
72	Machine Operators, Assorted Materials
73	Fabricators, Assemblers, and Hand Working Occupations
74	Production Inspectors, Testers, Samplers, and Weighers
75	Motor Vehicle Operators
76	Rail Transportation Occupations
77	Water Transportation Occupations
78	Material Moving Equipment Operators
79	Helpers, Construction and Extractive Occupations
80	Freight, Stock, and Material Handlers

1. *Education*: defined, for both parents and children, as the highest level of education attained.

2. *Earnings*: defined, for both parents and children, as the average earnings in the most frequently held occupation between age 22 and 55. Our earnings measure reflects wages and salaries, inclusive of bonus payments. Prior to constructing earnings in the cross-section, in the panel we remove age and time trends by projecting earnings on a quadratic age term, a quadratic time trend and an interaction term between age and year, all of which are allowed to vary by occupation. We then evaluate earnings at age 40 and in year 2000, to ensure comparability across years when averaging over time. The earnings variable in the cross-section is then obtained by averaging over the earnings in the most frequently held occupation. Since the earnings variable thus constructed nets out age and time effects, in all subsequent regressions we do not control

for age and time. Although we do not explicitly control for cohort fixed effects, we verify ex-post that the earnings variable is relatively stable across cohorts of parents and children.

We make a few additional remarks that apply to this, as well as other variables in the analysis. First, earnings, as well as all other nominal variables used in the analysis are expressed in 1996 US dollars. Second, earnings of the parent refer to the sum between the earnings of the father and the earnings of the mother. Third, the parent's age, occupation, and education refer to those pertaining to the head of the parent household, which is usually the father.

3. *Parental income*: defined as the average of the parent's family income between age 22 and 55. Our income measure equals the sum of taxable income, transfers and social security income of all members of the family unit. As with earnings, we first remove age and time trends by projecting family income on a quadratic age term, a quadratic time trend and an interaction term between age and year, all of which are allowed to vary by occupation. We allow these to vary by occupation as labor earnings is a component of family income. We then evaluate family income at the age of 40 and in year 2000, to ensure comparability across years when averaging over time, and do not control for age or time in any subsequent regression that uses this variable. Although we do not explicitly control for cohort fixed effects, we verify ex-post that the parental income variable is relatively stable across cohorts of parents.

4. *Parental endowment*: defined as the sum between parental earnings and annualized parental inherited wealth. Parental earnings is constructed as described above. As for parental inherited wealth, PSID only collected information on household wealth in 1984, 1989, 1994 and every other year since 1999. To bypass this data limitation we pursue the following imputation procedure. Let a_{it} denote the wealth household i in year t, and x_{it} denote a vector of observable characteristics of household i in year t that includes earnings, family income, full sets of dummies for age, race, family size, marital status, years of schooling and calendar year. We first estimate the following cross-validation lasso model

$$\min_{\theta} \sum \left(a_{it} - x_{it}'\theta\right)^2 + \lambda \left\|\theta\right\|_1 ,$$

where θ is a vector of parameters and λ is the penalty level, both to be estimated. The penalty level λ is chosen by cross-validation in order to optimize out-of-sample prediction performance. We consider a 5-fold cross-validation, which means that the

the data is split into 5 parts and the estimator is trained on all but the k^{th} fold and then validated on the k^{th} fold, iterating over $k = 1, \ldots, 5$. We then use the estimate of θ, which we denoted by $\hat{\theta}$,to impute wealth, when missing, according to$\hat{a}_{it} = x'_{it}\hat{\theta}$. We note that for the observations with non-missing wealth, projecting observed wealth a_{it} on imputed wealth $\hat{a}_{it}$ yields a slope of 1.135 with a standard error of 0.009 and an R^2 of 0.31.

We define wealth in the cross-section as the average of parent's wealth between age 22 and 55. As before, to ensure comparability across time, we first project wealth on a quadratic age term, a quadratic time trend and an interaction term between age and calendar year and evaluate wealth at age 40 and in year 2000.

Lastly, letting $\hat{a}_i$ denote parental wealth in the cross-section and $\hat{e}_i$ denote parental earnings in the cross-section, both constructed as discussed above, we defined parental endowment y_i as

$$y_i = \hat{e}_i + \frac{\hat{a}_i \times 0.638}{30},$$

where $\hat{a}_i$ is multiplied by a factor of 0.638 to account for the fact that approximately 63.8% of wealth is inherited (Gale and Scholz, 1994) and then divided by 30 to account for the fact that in the model a period is 30 years.

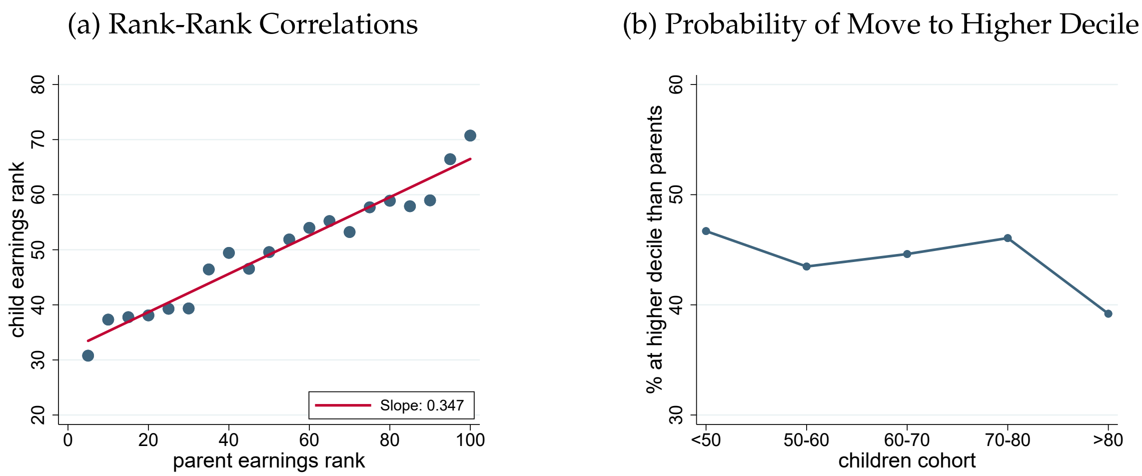
A.2 Intergenerational Mobility in PSID

We begin by revisiting the patterns of intergenerational mobility in the US using the PSID data. We compare our results with those reported by Chetty et al. (2014) using de-identified federal income tax records to establish that the PSID is indeed suitable for the study of intergenerational mobility.

The measure of intergenerational mobility we consider is in the tradition of Solon (1999), Dahl and DeLeire (2008), Black and Devereux (2011), Chetty et al. (2014), and reflects the relative outcomes of children from different parental backgrounds. The specific measure of relative mobility we employ is the *rank-rank slope*, the slope coefficient of a regression of the child's position in the earnings distribution on the position of their parent in the distribution. Parent and child earnings ranks are calculated relative to their corresponding birth cohort.

We estimate a rank-rank slope equal to 0.35, meaning that a 10 percentile point increase in parent's earnings rank is associated with a 3.5 percentile point increase in the child's earnings rank. Important to note is that the rank-rank slope estimated with the PSID data is almost identical to the value 0.34 reported in Chetty et al. (2014) based on administra-

Figure 14: Intergenerational Mobility of Earnings in PSID Data



Notes: Panel (a) plots the mean child rank within each parent earnings bin. There are 20 bins. Panel (b) displays the fraction of children born in the cohort on the X-axis who are in a higher decile of the lifetime earnings distribution than their parents.

tive data. This suggests that the PSID is representative of the US population in terms of intergenerational mobility and is thus suited for the analysis in this paper.

We also calculate a measure of absolute intergenerational mobility, namely the fraction of children who move to a higher earnings decile than their parents. On average, 43% of children move to a higher decile of the lifetime earnings distribution than their parents. This fraction is however declining over time, consistent with the findings of Chetty et al. (2014) and Chetty et al. (2017). Figure 14a and Figure 14b below offer a depiction of these statistics. Figure 14a displays the mean child earnings rank for 20 parent earnings bins. Figure 14b displays the evolution of the fraction of children who move to a higher earnings decile than their parents for five birth cohorts of children: pre-1950, 1951-1960, 1961-1970, 1971-1980, post-1980.

A.3 NLSY Data and Sample Selection

The NLSY is a longitudinal survey of a nationally representative sample of approximately 9,000 youths who were between 12 and 16 years old as of December 31, 1996. The first round of interviews took place in 1997, when both the youths and their parents were interviewed. In subsequent years, the youths were interviewed annually until 2011 and biennially since then. We use the NLSY to complement our PSID analysis of occupational choice as a function of parental income. As with the PSID, we transform the panel into a cross-section with information on the occupation, education and earnings of the children, as well as the lifetime

income of parents.

We apply the same procedure as with the PSID for transforming the panel data into a cross-section. Specifically, we define the occupation of the child as the most frequently held occupation between age 22 and age 36, the maximum age in the NLSY sample. We define education as the highest level of education attained and labor earnings as the average earnings in the most frequently held occupation between age 22 and 36, net of age and time effects that are allowed to vary by occupation. Between 1997 and 2003 the survey collected information on the income of the parent. We we define parental income in the cross-section as the average over parental family income over this period, net of time effects.

We make use of all the waves of the NLSY 1997. We transform the panel into a cross-section following, as closely as possible, the procedure applied to the PSID data. The result cross-section contains the following variables:

1. *Occupation*: defined as the most frequenty held occupation between age 22 and age 36. The oldest respondents in the NLSY 1997 are 36. We use the occupation classification in Table 5.

2. *Education*: defined as the highest level of education attained.

3. *Earnings*: defined as the average earnings in the most frequently held occupation between age 22 and 36. Prior to constructing earnings in the cross-section, in the panel we remove age and time trends by projecting earnings on a quadratic age term, a quadratic time trend and an interaction term between age and year, all of which are allowed to vary by occupation. We then evaluate earnings at age 30 and in year 2010, to ensure comparability across years when averaging over time. We evaluate earnings at a different age and in a different year than in the PSID data because the NLSY sample covers a more recent period than the PSID. The earnings variable in the cross-section is then obtained by averaging over the earnings in the most frequently held occupation. Since the earnings variable thus constructed nets out age and time effects, in all regressions that use this variable we do not control for age and time.

4. *Parental income*: defined as the average of the parent's family income collected in the survey. We first remove time trends by projecting parental income on a quadratic time trend. We then evaluate family income in year 2010.

A.4 General Social Survey

The GSS is a survey that assesses attitudes, behaviors, and attributes of a representative sample of US residents. The survey began in 1972, collecting information on a sample between

1,500 and 4,000 respondents. We use seven topics/questions from the Quality of Working Life module, administered in 2002, 2006, 2010 and 2014. These topics/questions are: *(i)* At the place where I work, I am treated with respect, *(ii)* Does your job regularly require you to perform repetitive or forceful hand movements or involved awkward postures?, *(iii)* Does your job require you to do repeated lifting, pushing, pulling or bending?, *(iv)* My job requires that I keep learning new things, *(v)* I have an opportunity to develop my own special abilities, *(vi)* I get to do a number of different things on my job, and *(vii)* My job requires that I work very fast. We recode answers to topics/questions *(i)*, *(iv)*-*(vii)* to range from 1-Strongly disagree, 2-Disagree, 3-Agree to 4-Strongly agree and answers to topics/questions *(ii)* and *(iii)* to 1-Yes and 2-No. We standardize these answers before estimating Equation 1, so that each v_{it}^{x} is a number between 0 and 1.

A.5 Empirical Analysis

A.5.1 Estimating Potential Earnings

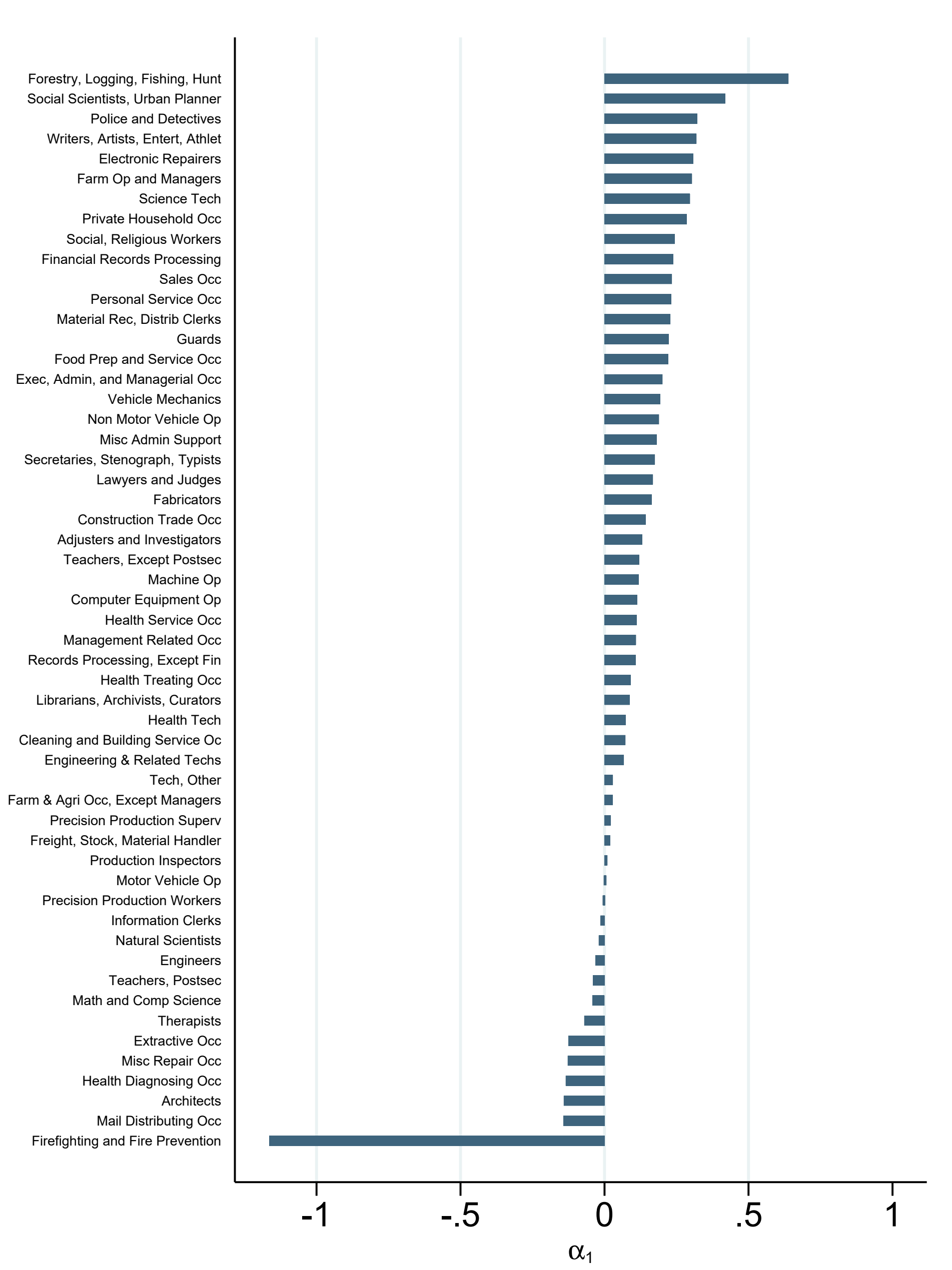
We examine the extent to which parental income increasses the efficiency of children in different occupations by estimating the following specification

$$\ln e_{ij} = \alpha_{1j} \ln \bar{y}_i + \tilde{\boldsymbol{X}}_i' \boldsymbol{\alpha}_j + \delta_j + \epsilon_{ij},$$

where e_{ij} are the annual earnings of child i working in occupation j, $\bar{y}_i$ is their parent's lifetime income, $\tilde{\boldsymbol{X}}_{\boldsymbol{i}}$ is a vector of covariates including years of schooling, age, gender and race whose effect on earnings is allowed to vary by occupation, and δ_j are occupation fixed effects. The coefficients of interest are α_{1j}, which capture the effect of parental income on occupational efficiency.

Figure 15 displays the estimates of α_{1j} for the 54 occupations we consider. The elasticity of potential earnings with respect to parental income is positive for most occupations. However, a visual inspection of Figure 1a and Figure 15 reveals a mixed relationship between this elasticity and the intrinsic quality of occupations. For example, children are more likely to earn more as social scientists, urban planners, writers, artists, entertainers or athletes, all occupations with a relatively high intrinsic quality, if they have richer parents. At the same time, children with rich parents are less likely to earn high earnings as architects or teachers, also occupations with a relatively high intrinsic quality. More formally, the correlation between α_{1j}, the elasticity of earnings with respect to parental income, and v_j, the intrinsic quality of occupations, is small (-0.047) and not statistically significant (SE=0.139).

Figure 15: Effect of Parental Income on Child's Earnings



Notes: Bars are elasticities capturing the effect of parental income on earnings across occupations.

B Model Appendix

B.1 The Stationary Distribution of Endowment and Intergenerational Mobility

Assume that the earnings function $e_j(s,\cdot,y)$ is monotonically increasing in talent $u$for all occupations and define a corresponding inverse of the earnings function $\widetilde{E}_j^{-1}(\cdot;s,y)$ as:

$$u \equiv \widetilde{E}_j^{-1}(e_j(s,u,y);s,y).$$

We can write the cumulative distribution function for the earnings of the children of parents with endowment y as

$$F_e\left(e|y\right) = \mathbb{E}_s\left[F_e(e|s,y)\,\middle|\,h^*(y)\right], \tag{24}$$

where we have defined the cumulative distribution function $F_e(e|s,y)$ of the earnings of children with schooling s and parental endowment y as

$$F_e(e|s,y) \equiv \sum_{j=1}^{J}\int^{\widetilde{E}_j^{-1}(e;s,y)} \mu_j\left(s,u,y\right)\, d\mathbb{P}_u(u), \tag{25}$$

where the conditional occupational choice function μ_j satisfies Equation (9). Equation (25) accounts for two distinc effects of the parental endowment on child earnings discussed above. The first effect, that higher parental endowment may raise the earnings within the occupation, is reflected in the upper bound on the integral. The second effect, that parental endowment shapes the patterns of occupational choice, is reflected through the dependence of the term μ_j on parental endowment. Finally, Equation (24) accounts for the effect of parental investment in human capital on the distribution of earnings.

Given the conditional distribution of earnings, it is easy to see that the stationary cumulative distribution function of total endowment y has to satisfy the following fixed point condition

$$F_y\left(y^+\right) = \int_0^\infty F_e(y^+ - b^*(y)|y)\, dF_y(y), \tag{26}$$

where the conditional distribution of earnings $F_e(\cdot|y)$ satisfies Equations (24) and (25). The dispersion in total endowment is shaped by two distinct forces: the dependence of child earnings on parental total income $F_e(\cdot|y)$ as well as the direct parental transfer policy $b^*(y)$.

Mobility of Welfare We can further characterize the dependence of the welfare of the child on the parental endowment as

$$F_v\left(v^+|y\right) = \mathbb{E}_{s,u}\left[F_v\left(v^+|s,u,y\right)\,|\,h^*(y)\right], \tag{27}$$

where $F_v\left(v^+|y\right)$ satisfies Equation (18). Equation (27) additionally accounts for the contribution of parental endowment to the welfare of the children through its effect on schooling attainment. Conditional on attained schooling, Equation (18) shows the welfare effect of parental endowment through its direct effect on earnings and its indirect effect on the patterns of occupation choice. Finally, the long-run stationary distribution of welfare in the

model immediately follows from Equation (27) as $F_v(v) \equiv \int F_v(v|y)\, dF_y(y)$.

B.2 Proofs and Derivations

Sequential Formulation of the Problem of Generation t The problem laid out in Section 3.1.1 corresponds to the recursive formulation of the the following sequential problem faced by each generation t:

$$\max_{(c_{t'}, j_{t'}, b_{t'+1}, h_{t'+1})_{t'=t}^{\infty}} \mathbb{E}_t \left[\sum_{t'=t}^{\infty} \beta^{\tau-t} \left(\log c_\tau + \zeta v_{j_\tau} + \epsilon_{j_\tau \tau} \right) \right],$$

$$y_{t'} \geq c_{t'} + \frac{b_{t'+1}}{1+r_{t'}} + \varphi_{t'}(h_{t'+1}), \qquad t' \geq t,$$

$$y_{t'} = b_{t'} + e_{j_{t'} t'} \left(s_{t'}, u_{t'}, y_{t'-1} \right),$$

facing a sequence of *i.i.d.* shocks $(\boldsymbol{\epsilon}_{t'})_{t'=t}^{\infty}$, s_t, and u_t. The timing of the decisions are such that agents in period t choose their own occupation and consumption j_t and c_t, and the investments b_{t+1} and h_{t+1} given the histories of the outcomes of their dynastic line. However, as the recursive formulation above shows, the relevant aspect of their ancestoral history can be captured by the total income of their parents y_{t-1} (and the corresponding investment decisions b_t and h_t).

Lemma 1. *The expected utility of children in Equation (4) is given by Equation (6).*

Proof. Let $\boldsymbol{\epsilon} \equiv (\epsilon_j)_{j=1}^{J}$ be a tuple of *i.i.d.* random variables distributed according to a zero mean, with the cumulative distribution function

$$F(x) \equiv \mathbb{P}(\epsilon_j \leq x) = \prod_{j=1}^{J} \exp(-\exp(-x - \overline{\gamma})),$$

where $\overline{\gamma} \equiv \int_{-\infty}^{\infty} u \exp(-u \exp(-u))du$ is the Euler-Mascheroni constant. Consider a child with schooling s, talent u, parental transfer b, and parental income y, and let $\vartheta_j \equiv V\left(b + e_j\left(s, u, y\right)\right) + \zeta\, v_j$,to simplify the expressions. The probability that the expected adult utility of this child is below v is given by

$$F_v(v) \equiv \mathbb{P}\left[V^{+}\left(s, u, \boldsymbol{\epsilon}, b, y\right) < v\right],$$

$$= \mathbb{P}\left[\max_j \ \vartheta_j + \rho\, \epsilon_j < v\right],$$

$$
\begin{aligned}
&= \prod_{j=1}^{J} \mathbb{P}\left(\epsilon_j \le \tfrac{1}{\rho}\left(v-\vartheta_j\right)\right), \\
&= \prod_{j=1}^{J} F\left(\tfrac{1}{\rho}\left(v-\vartheta_j\right)\right), \\
&= \prod_{j=1}^{J} \exp\left(-\exp\left(-\tfrac{1}{\rho}\left(v-\vartheta_j\right)-\overline{\gamma}\right)\right), \\
&= \exp\left(-\exp\left[-\tfrac{1}{\rho}v+\log\left(\sum_{j=1}^{J} e^{\frac{1}{\rho}\vartheta_j}\right)-\overline{\gamma}\right]\right).
\end{aligned}
$$

This allows us to calculate

$$
\begin{aligned}
\mathbb{E}_\epsilon\left[V^+\left(s,u,\epsilon,b,y\right)\right] &= \tfrac{1}{\rho}\sum_{j=1}^{J}\int_{-\infty}^{\infty} v\, f\left(\tfrac{1}{\rho}\left(v-\vartheta_j\right)\right)\prod_{j'\neq j} F\left(\tfrac{1}{\rho}\left(v-\vartheta_j\right)\right) dv, \\
&= \tfrac{1}{\rho}\sum_{j=1}^{J}\int_{-\infty}^{\infty} v\, e^{-\frac{1}{\rho}\left(v-\vartheta_j\right)-\overline{\gamma}} \prod_{j'=1}^{J}\exp\left(-\exp(-\tfrac{1}{\rho}\left(v-\vartheta_{j'}\right)-\overline{\gamma})\right) dv, \\
&= \tfrac{1}{\rho}\int_{-\infty}^{\infty} v\left(e^{-\frac{1}{\rho}v-\overline{\gamma}}\sum_{j=1}^{J} e^{\frac{1}{\rho}\vartheta_j}\right)\exp\left(e^{-\frac{1}{\rho}v-\overline{\gamma}}\sum_{j'=1}^{J} e^{\frac{1}{\rho}\vartheta_{j'}}\right) dv.
\end{aligned}
$$

Defining $x \equiv \frac{1}{\rho}v+\overline{\gamma}-\log\sum_{j'=1}^{J} e^{\frac{1}{\rho}\vartheta_{j'}}$, we find:

$$
\begin{aligned}
\mathbb{E}_\epsilon\left[V^+\left(s,u,\epsilon,b,y\right)\right] &= \rho\sum_{j=1}^{J}\int_{-\infty}^{\infty}\left(x-\overline{\gamma}+\log\sum_{j'=1}^{J}\exp\left(\tfrac{1}{\rho}\vartheta_{j'}\right)\right)\exp\left(-x\right)\exp\left(\exp\left(-x\right)\right)dx, \\
&= \rho\log\sum_{j'=1}^{J}\exp\left(\tfrac{1}{\rho}\vartheta_{j'}\right).
\end{aligned}
$$

□

Lemma 2. *The probabilities of occupational choice under a stationary distribution is given by Equation (9).*

Proof. We use the same notation as in the proof of Lemma 1 above. Dropping the time subscripts to simplify the expressions, the probably of choosing occupation j for a child with schooling s, talent u, parental transfer b, and parental income y is given by

$$
\mu_j\left(s,u,b,y\right) \equiv \mathbb{P}\left(j = \underset{j'}{\operatorname{argmax}}\ \vartheta_{j'}+\rho\,\epsilon_{j'}\right),
$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} F'\left(\epsilon_j\right) \times \prod_{j' \neq j} \mathbb{P}\left(\epsilon_{j'} \leq \epsilon_j + \tfrac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)\right) d\epsilon_j, \\
&= \int_{-\infty}^{\infty} \exp\left(-\epsilon_j - \overline{\gamma}\right) \exp\left(-e^{-\epsilon_j - \overline{\gamma}}\right) \\
&\quad \times \prod_{j' \neq j} \exp\left(-e^{-\left(\epsilon_j + \frac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)\right) - \overline{\gamma}}\right) d\epsilon_j, \\
&= \int_{-\infty}^{\infty} \exp\left(-\epsilon_j - \overline{\gamma}\right) \exp\left(-e^{-\epsilon_j - \overline{\gamma}}\left(1 + \sum_{j' \neq j} e^{-\frac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)}\right)\right) d\epsilon_j, \\
&= \frac{1}{1 + \sum_{j' \neq j} e^{-\frac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)}} \int_0^{\infty} \exp\left(-x\right) dx, \\
&= \frac{e^{\frac{1}{\rho}\vartheta_j}}{\sum_{j'} e^{\frac{1}{\rho}\vartheta_{j'}}},
\end{aligned}
$$

where in the last equality, we have used the change of variables $x \equiv e^{-\epsilon_j - \overline{\gamma}}\left(1 + \sum_{j' \neq j} e^{-\frac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)}\right)$. $\square$

Lemma 3. *For a stationary equilibrium, define V^+ as in Equation (2). We then have $\mathbb{P}\left(V^+ < v | y, s, u, j\right) = \mathbb{P}\left(V^+ < v | y, s, u\right)$, where we have defined the distribution of utility conditional on the selected occupation j as*

$$
\mathbb{P}\left(V^+ < v | y, s, u, j\right) \equiv \mathbb{P}\left(V^+ < v \,\middle|\, y, s, u,\, j = \underset{j'}{\mathit{argmax}}\, V\left(b + e_{j'}\left(s, u, y\right)\right) + \zeta\, \nu_{j'} + \rho\, \epsilon_{j'}\right).
$$

Proof. We use the same notation as in the proof of Lemma 2 above. The distribution of utilities, conditional on a given occupation j is given by:

$$
\begin{aligned}
F_v\left(v | j\right) &\equiv \mathbb{P}\left(V^+\left(s, u, \boldsymbol{\epsilon}, b, y\right) < v |\, j = \underset{j'}{\operatorname{argmax}}\, \vartheta_{j'} + \rho\, \epsilon_{j'}\right), \\
&= \frac{\mathbb{P}\left(V^+\left(s, u, \boldsymbol{\epsilon}, b, y\right) < v\,,\, j = \underset{j'}{\operatorname{argmax}}\, \vartheta_{j'} + \rho\, \epsilon_{j'}\right)}{\mathbb{P}\left(j = \underset{j'}{\operatorname{argmax}}\, \vartheta_{j'} + \rho\, \epsilon_{j'}\right)}, \\
&= \frac{1}{\mu_j} \times \int_{-\infty}^{\frac{1}{\rho}\left(v - \vartheta_{jt}\right)} F'\left(\epsilon_j\right) \times \prod_{j' \neq j} \mathbb{P}\left(\epsilon_{j'} \leq \epsilon_j + \tfrac{1}{\rho}\left(\vartheta_j - \vartheta_{j'}\right)\right) d\epsilon_j,
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\mu_j}\int_{-\infty}^{\frac{1}{\rho}(v-\vartheta_{jt})} \exp\left(-\epsilon_j-\overline{\gamma}\right)\exp\left(-e^{-\epsilon_j-\overline{\gamma}}\left(1+\sum_{j'\neq j} e^{-\frac{1}{\rho}\left(\vartheta_j-\vartheta_{j'}\right)}\right)\right)d\epsilon_j,\\
&= \frac{1}{\mu_j}\times\frac{1}{1+\sum_{j'\neq j} e^{-\frac{1}{\rho}\left(\vartheta_j-\vartheta_{j'}\right)}}\int_{e^{-\frac{1}{\rho}\left(v-\vartheta_{jt}\right)-\overline{\gamma}}\left(1+\sum_{j'\neq j} e^{-\frac{1}{\rho}\left(\vartheta_j-\vartheta_{j'}\right)}\right)}^{\infty}\exp\left(-x\right)dx,\\
&= \exp\left(-e^{-\frac{1}{\rho}v-\overline{\gamma}}\left(\sum_j e^{\frac{1}{\rho}\vartheta_j}\right)\right),\\
&= F_v(v),
\end{aligned}
$$

where, again, in the fifth equality we have used the change of variables

$$
x \equiv e^{-\epsilon_j-\overline{\gamma}}\left(1+\sum_{j'\neq j} e^{-\frac{1}{\rho}\left(\vartheta_j-\vartheta_{j'}\right)}\right).
$$

□

Lemma 4. *The joint distribution of the observed data based on the model is given by*

$$
\begin{aligned}
\mathbb{P}\left(\boldsymbol{d};\varsigma\right) = \prod_{i=1}^{N}\Bigg\{&\frac{\exp\left[\frac{\zeta}{\rho}\nu_{o_i}+\frac{1}{\rho}V\left(b^*\left(y_i\right)+e_{o_i}\left(s_i,\mathcal{U}\left(e_i,s_i,o_i,y_i;\varsigma\right),y_i\right)\right)\right]}{\sum_j\exp\left[\frac{\zeta}{\rho}\nu_j+\frac{1}{\rho}V\left(b^*\left(y_i\right)+e_j\left(s,\mathcal{U}\left(e_i,s_i,o_i,y_i;\varsigma\right),y\right)\right)\right]}\\
&\times\frac{1}{\sqrt{2\pi\,\theta_{o_i}^2}}\exp\left(-\frac{1}{2}\mathcal{U}\left(e_i,s_i,o_i,y_i;\varsigma\right)^2\right)\times\frac{\exp\left(-\frac{1}{2}\left(\frac{s_i-h^*(y_i)}{\vartheta}\right)^2\right)}{\sum_{s'=0}^{4}\exp\left(-\frac{1}{2}\left(\frac{s'-h^*(y_i)}{\vartheta}\right)^2\right)}\Bigg\},
\end{aligned}
\tag{28}
$$

where $\mathcal{U}\left(e_i,s_i,o_i,y_i;\varsigma\right)$ is defined by Equation (16).

Proof. The observations are independent, thus we have $\mathbb{P}\left(\boldsymbol{d};\varsigma\right)=\prod_i\mathbb{P}\left(e_i,o_i,s_i|y_i\right)$. Based on the model, we have:

$$
\begin{aligned}
\mathbb{P}\left(e_i,o_i,s_i|y_i\right) &= \mathbb{E}_{u_i}\left[\mathbb{P}\left(e_i,o_i,y_i,s_i,u_i\right)\right],\\
&= \int\mathbb{P}\left(o_i|y_i,s_i,u_i\right)\delta\left(e_i-\left(\alpha_{o_i}+\kappa_{o_i}s_i+\delta_{o_i}y_i+\theta_{o_i}u_i\right)\right)\mathbb{P}\left(u_i\right)\mathbb{P}\left(s_i|y_i\right)du_i,\\
&= \mathbb{P}\left(s_i|y_i\right)\int\mathbb{P}\left(o_i|y_i,s_i,u_i\right)\delta\left(e_i-\left(\alpha_{o_i}+\kappa_{o_i}s_i+\delta_{o_i}y_i+\theta_{o_i}u_i\right)\right)\frac{e^{-u_i^2/2}}{\sqrt{2\pi}}du_i,\\
&= \mathbb{P}\left(s_i|y_i\right)\int\mathbb{P}\left(o_i|y_i,s_i,\frac{x}{\theta_{o_i}}\right)\delta\left(e_i-\left(\alpha_{o_i}+\kappa_{o_i}s_i+\delta_{o_i}y_i\right)-x\right)\frac{e^{-x^2/2\theta_{o_i}^2}}{\sqrt{2\pi}}\frac{dx}{\theta_{o_i}},
\end{aligned}
$$

$$= \mathbb{P}_{s|h}\left(s_i|h^*\left(y_i\right)\right)\mathbb{P}\left(o_i|y_i, s_i, \frac{e_i - \left(\alpha_{o_i} + \kappa_{o_i}s_i + \delta y_i\right)}{\theta_{o_i}}\right)\frac{e^{-\left(e_i - \left(\alpha_{o_i} + \kappa_{o_i}s_i + \delta y_i\right)\right)^2/2\theta_{o_i}^2}}{\sqrt{2\pi}\theta_{o_i}},$$

where we have perform the change of variables $x \equiv u_i/\theta_{o_i}$ in the fourth equality. Equation (28) immediately follows. □

C Estimation Appendix

C.1 Log-Likelihood Function

The maximum-likelihood estimation problem corresponds to that of maximizing the following the log-likelihood function

$$\begin{aligned}
\mathcal{L}\left(\boldsymbol{d};\varsigma\right) \equiv & \sum_{i=1}^{N}\log\mathbb{P}\left(e_i, o_i, s_i|y_i\right), \\
= & \frac{\zeta}{\rho}\left(\sum_{i=1}^{N}\nu_{o_i}\right) + \frac{1}{\rho}\sum_{i=1}^{N}V\left(b^*\left(y_i\right) + e_j\left(s_i, y_i, \mathcal{U}\left(e_i, o_i, s_i, y_i;\varsigma\right)\right)\right) \\
& - \sum_i\log\left(\sum_j\exp\left[\frac{\zeta}{\rho}\nu_j + \frac{1}{\rho}V\left(b^*\left(y_i\right) + e_j\left(s_i, y_i, \mathcal{U}\left(e_i, o_i, s_i, y_i;\varsigma\right)\right)\right)\right]\right) \\
& - \frac{1}{2}\sum_i\mathcal{U}\left(e_i, o_i;\varsigma\right)^2 - \tfrac{1}{2}\sum_i\log\theta_{o_i} \\
& - \frac{1}{2}\sum_i\left(\frac{s_i - h^*\left(y_i\right)}{\vartheta}\right)^2 - \sum_i\log\sum_{s'=0}^{4}\exp\left(-\frac{1}{2}\left(\frac{s' - h^*\left(y_i\right)}{\vartheta}\right)^2\right).
\end{aligned} \tag{29}$$

The second and third lines of Equation (29) characterize the conditional distribution of occupational choice, given schooling, earnings, and parental endowment. The fourth and fifth lines characterize the conditional distribution of talent and schooling, given parental endowment. We find the set of parameters ς maximizing the log likelihood function above for the observed data. For the derivation, see Lemma 4 in Appendix B.

C.2 Details of the Estimation Procedure

Initialization. We initialize the values of parameters in our main estimation based on a preliminary estimation stage using a less granular classification of the occupations observed in the data at the level of 14 occupation codes. We further simplify the parameter search in this initial stage by setting the return to parental endowments in occupation-specific ability

to zero, i.e., $\boldsymbol{\delta} \equiv \mathbf{0}$.

In turn, we initialize the values of parameters of the restricted model in this first stage estimation by applying the following strategy. First, note that Equations (14) and (9) together imply that the conditional expected log earnings of children based on the model satisfies:

$$\mathbb{E}\left[\log\left(e\right)|j,s,y\right] = \alpha_j + \kappa_j \log s + \delta_j \log y + \theta_j \underbrace{\frac{\int u \times \mu_j\left(s,u,y\right) d\mathbb{P}(u)}{\int \mu_j\left(s,u,y\right) d\mathbb{P}(u)}}_{\equiv \overline{u}_j(s,y)}, \tag{30}$$

where $\overline{u}_j\left(s,y\right)$ stands for the conditional expectation of talent given parental endowment, schooling, and occupational choice. This term controls for the effect of selection on unobservable talent and shows why we cannot uncover the occupation-specific returns to schooling and parental endowment based on a simple regression of log earnings on the latter. We can similarly derive the conditional variance of log earnings as:

$$\mathbb{V}\left[\log\left(e\right)|j,s,y\right] = \theta_j^2 \frac{\int \left(u - \overline{u}_j\left(s,y\right)\right)^2 \times \mu_j\left(s,u,y\right) d\mathbb{P}(u)}{\int \mu_j\left(s,u,y\right) d\mathbb{P}(u)}. \tag{31}$$

Intuitively, the presence of a strong dispersion in log earnings in a given occupation conditional on schooling and parental income suggests a strong degree of return to talent in that occupation.

We consider a set of bins for the values of parental endowment $Y = \{\bar{y}^1, \bar{y}^2, \bar{y}^3, \bar{y}^4, \bar{y}^5\}$ and map each observed parental endowment in the data to one of the bins, setting $\bar{y}_i \equiv \arg\min_{\bar{y}\in Y} \left| \log y_i - \bar{y} \right|$. Inspired by Equations (30) and (31), we find an initial estimate for the coordinates of returns to schooling $\boldsymbol{\kappa}$ by relying on an observation-weighted least-squares regression of log earnings $\widehat{\mathbb{E}}\left[\log e \middle| j,s,\bar{y}\right]$ on schooling s while attempting to control for the selection term by $\widehat{\mathbb{V}}\left[\log e \middle| j,s,\bar{y}\right]^{1/2}$. Using the resulting estimates, we recover initial guesses for occupation-specific fixed earnings and returns to talent $(\boldsymbol{\alpha}, \boldsymbol{\theta})$ as

$$\alpha_j^{(0)} = \frac{\sum_{s,\bar{y}} \left(\widehat{\mathbb{E}}\left[\log e \middle| j,s,\bar{y}\right] - \kappa s\right) \#(j,s,\bar{y})}{\sum_{s,\bar{y}} \#(j,s,\bar{y})},$$
$$\theta_j^{(0)} = \sqrt{\frac{\sum_{s,\bar{y}} \widehat{\mathbb{V}}\left[\log e \middle| j,s,\bar{y}\right] \#(j,s,\bar{y})}{\sum_{s,\bar{y}} \#(j,s,\bar{y})}}.$$

The procedure above yields our initial guesses for the parameters of the earnings function. For the remaining parameters, we pick the following initial guesses. In practice, we pa-

rameterize the cost function $\varphi(\cdot)$ for human capital investments with a vector $(\widetilde{\varphi}_1, \widetilde{\varphi}_2, \widetilde{\varphi}_3, \widetilde{\varphi}_4)$ such that $\varphi_k \equiv \exp(\widetilde{\varphi}_k)$ gives the slope of the cost function in the interval $h \in [k-1, k]$. We consider a convex form characterized by $\widetilde{\boldsymbol{\phi}} = (5, 6, 7, 8)$. Finally, we initialize the remaining parameters, i.e., the dispersion of idiosyncratic taste shocks ρ, the weight of intrinsic valuations ζ, and the dispersion of schooling shocks ϑ all at unity.

Optimization. We perform the maximization of the log likelihood objective function in two stages. In the first stage, we perform an iterative, block-wise scheme, in which we iterate over maximizing the objective function only over one of the following three partitions of the model parameters (keeping all other components at their current levels): 1) the taste parameters (ζ, ρ), 2) the human capital cost parameters $(\widetilde{\boldsymbol{\phi}}, \vartheta)$,[44] and 3) the parameters of the earnings function $(\boldsymbol{\alpha}, \boldsymbol{\kappa}, \boldsymbol{\theta}, \boldsymbol{\delta})$.[45] After a few rounds of this block-wise optimization, we then perform a joint maximization of the objective function over the entire parameter space using a SQP-type algorithm.

54-Occupations Environment. We use the estimates found on the data with 14 occupational codes to initialize the parameters of the model for the main data with 54 occupational codes. We rely on a crosswalk between the two levels to initialize all the parameters of the earnings function at the 54-occupation level that belong to the same 14-occupation level code with the values estimated in the first stage for the latter. We then apply another iterative, block-wise optimization scheme similar to the one discussed above across the implied 14 blocks of occupational codes. For each block, we separately update the parameters of the earnings function corresponding to the occupations within each of the 14-occupation codes. After a few rounds of applying this block-wise strategy, we follow the same strategy as that discussed above for the 14-code level to gradually extend the search to the joint space including other model parameters. We finally introduce the returns to parental endowment parameters $\boldsymbol{\delta}$, before applying a final round of joint optimization in the space of all model parameters.[46]

[44]In practice, we found overall improvements in the final objective function when in the rounds updating the education parameter block we initially over-weight the terms in the objective function that correspond to the conditional distribution of schooling attainment given parental endowments.

[45]Since we rely on a discretization of the state space to solve the Bellman equation to compute the objective function, the numerical evaluation of the gradients and the Jacobians of the objective function often leads to discontinuities. In order to smooth out these discontinuities, we streer the optimization routine by providing initially large-step approximations to the gradients and gradually lowering the step-size for the evaluation of the gradients.

[46]We initialize the values of these returns parameters as the slopes corresponding to auxiliary regressions of $\theta_j u_i$ on y_i for all i such that $o_i = j$.

C.3 Estimated Earnings Function

Table 7 reports the estimated parameters of the earnings function for each occupation.

Table 7: Estimated Earnings Function

Occ	Description	α	κ	θ	δ
1	Executive, Administrative, and Managerial Occupations	7.876	0.203	0.188	0.540
		(0.081)	(0.021)	(0.007)	(0.077)
2	Management Related Occupations	7.704	0.233	0.190	0.549
		(0.078)	(0.023)	(0.007)	(0.080)
3	Architects	7.478	0.241	0.188	0.579
		(0.052)	(0.066)	(0.016)	(0.098)
4	Engineers	7.515	0.251	0.194	0.578
		(0.054)	(0.026)	(0.009)	(0.085)
5	Mathematical and Computer Scientists	7.735	0.194	0.193	0.568
		(0.083)	(0.031)	(0.008)	(0.092)
6	Natural Scientists	7.458	0.272	0.188	0.586
		(0.080)	(0.045)	(0.012)	(0.089)
7	Health Diagnosing Occupations	7.408	0.303	0.183	0.608
		(0.091)	(0.076)	(0.022)	(0.102)
8	Health Assessment and Treating Occupations	7.728	0.244	0.184	0.548
		(0.059)	(0.023)	(0.007)	(0.084)
9	Therapists	7.468	0.281	0.188	0.560
		(0.057)	(0.032)	(0.011)	(0.097)
10	Teachers, Postsecondary	7.352	0.309	0.185	0.561
		(0.076)	(0.059)	(0.014)	(0.093)
11	Teachers, Except Postsecondary	7.428	0.292	0.198	0.497
		(0.079)	(0.017)	(0.010)	(0.046)
12	Librarians, Archivists, and Curators	7.358	0.267	0.189	0.488
		(0.052)	(0.059)	(0.011)	(0.121)
13	Social Scientists and Urban Planners	7.451	0.276	0.186	0.597
		(0.065)	(0.046)	(0.015)	(0.095)
14	Social, Recreation, and Religious Workers	7.428	0.266	0.199	0.473
		(0.059)	(0.022)	(0.009)	(0.074)
15	Lawyers and Judges	7.470	0.283	0.184	0.599
		(0.081)	(0.047)	(0.014)	(0.091)

16	Writers, Artists, Entertainers, and Athletes	7.534	0.233	0.202	0.528
		(0.088)	(0.014)	(0.008)	(0.071)
17	Health Technologists and Technicians	7.806	0.178	0.198	0.486
		(0.060)	(0.028)	(0.009)	(0.100)
18	Engineering and Related Technologists and Technicians	7.811	0.156	0.192	0.566
		(0.076)	(0.042)	(0.015)	(0.094)
19	Science Technicians	7.710	0.157	0.193	0.554
		(0.066)	(0.049)	(0.012)	(0.111)
20	Technicians, Except Health, Engineering, and Science	7.676	0.243	0.183	0.589
		(0.060)	(0.049)	(0.014)	(0.097)
21	Sales Occupations	7.953	0.175	0.192	0.507
		(0.086)	(0.020)	(0.007)	(0.076)
22	Miscellaneous Administrative Support Occupations	7.886	0.156	0.195	0.468
		(0.087)	(0.016)	(0.009)	(0.068)
23	Computer and Communication Equipment Operators	7.863	0.066	0.202	0.489
		(0.096)	(0.076)	(0.019)	(0.099)
24	Secretaries, Stenographers, and Typists	7.909	0.151	0.195	0.457
		(0.068)	(0.015)	(0.006)	(0.069)
25	Information Clerks	7.789	0.147	0.201	0.458
		(0.091)	(0.036)	(0.010)	(0.088)
26	Records Processing Occupations, Except Financial	7.741	0.179	0.193	0.538
		(0.070)	(0.044)	(0.012)	(0.104)
27	Financial Records Processing Occupations	7.857	0.157	0.196	0.482
		(0.057)	(0.024)	(0.010)	(0.089)
28	Mail Distribution Occupations	7.890	0.092	0.202	0.537
		(0.092)	(0.063)	(0.019)	(0.098)
29	Material Recording, Scheduling, and Distributing Clerks	7.920	0.133	0.199	0.463
		(0.062)	(0.023)	(0.012)	(0.090)
30	Adjusters and Investigators	7.883	0.143	0.196	0.495
		(0.066)	(0.023)	(0.007)	(0.091)
31	Private Household Occupations	7.888	0.047	0.203	0.414
		(0.074)	(0.021)	(0.008)	(0.072)
32	Guards	7.816	0.144	0.195	0.535
		(0.088)	(0.041)	(0.015)	(0.098)
33	Firefighting and Fire Prevention Occupations	7.550	0.143	0.205	0.599

		(0.076)	(0.044)	(0.013)	(0.085)
34	Police and Detectives	7.785	0.189	0.191	0.559
		(0.084)	(0.029)	(0.010)	(0.091)
35	Food Preparation and Service Occupations	7.957	0.126	0.200	0.423
		(0.067)	(0.015)	(0.009)	(0.066)
36	Health Service Occupations	7.833	0.146	0.202	0.431
		(0.063)	(0.017)	(0.008)	(0.073)
37	Cleaning and Building Service Occupations	7.932	0.117	0.199	0.477
		(0.072)	(0.027)	(0.012)	(0.087)
38	Personal Service Occupations	7.877	0.147	0.195	0.483
		(0.085)	(0.014)	(0.008)	(0.062)
39	Farm Operators and Managers	7.561	0.205	0.200	0.537
		(0.073)	(0.028)	(0.011)	(0.075)
40	Farm and Agricultural Occupations, Except Managerial	7.868	0.048	0.207	0.464
		(0.082)	(0.029)	(0.015)	(0.084)
41	Forestry, Logging, Fishing and Hunting Occupations	7.539	0.134	0.220	0.573
		(0.051)	(0.019)	(0.010)	(0.058)
42	Vechicle Mechanics	7.931	0.120	0.200	0.486
		(0.075)	(0.019)	(0.010)	(0.092)
43	Electronic Repairers	7.861	0.129	0.196	0.547
		(0.070)	(0.031)	(0.009)	(0.087)
44	Miscellaneous Repair Occupations	7.796	0.113	0.197	0.584
		(0.099)	(0.034)	(0.013)	(0.092)
45	Construction Trade Occupations	7.953	0.114	0.201	0.498
		(0.137)	(0.020)	(0.015)	(0.089)
46	Extractive Occupations	7.520	0.098	0.220	0.594
		(0.074)	(0.022)	(0.011)	(0.062)
47	Precision Production Supervisors	7.771	0.137	0.199	0.555
		(0.074)	(0.016)	(0.007)	(0.081)
48	Precision Production Workers	7.879	0.123	0.199	0.501
		(0.079)	(0.025)	(0.010)	(0.090)
49	Machine Operators	8.041	0.092	0.198	0.463
		(0.078)	(0.023)	(0.011)	(0.089)
50	Fabricators	7.952	0.100	0.201	0.509
		(0.086)	(0.021)	(0.012)	(0.091)

51	Production Inspectors	7.797	0.139	0.196	0.551
		(0.073)	(0.047)	(0.013)	(0.092)
52	Motor Vehicle Operators	8.056	0.085	0.198	0.457
		(0.066)	(0.024)	(0.011)	(0.093)
53	Non Motor Vehicle Operators	7.959	0.097	0.200	0.475
		(0.104)	(0.023)	(0.011)	(0.086)
54	Freight, Stock and Material Handlers	7.953	0.070	0.204	0.500
		(0.108)	(0.033)	(0.015)	(0.074)

Notes: Table entries show the estimated parameters of the earnings function. Standard errors for each parameter, computed based on re-estimating the model for 25 bootstrapped samples, are in the parentheses.

C.4 Additional Estimation Results

C.4.1 Untargeted Moments

Table 8 compares the predictions of the model regarding children's schooling attainment as a function of parental endowment with the corresponding patterns in the data. Consistent with the data, children of poor parents (i.e. those with log parental endowment below the median) in the model are more likely not to graduate from high-school or to only obtain a high-school degree. Conversely, children of rich parents have a higher educational attainment and are more likely to obtain a college or a graduate degree.

Table 9 assesses the model's performance in terms of predicting the dependence of occupational choice on parental endowment and schooling attainment. To that end, the ta-

Table 8: Schooling Attainment Conditional on Parental Endowment

	Data		Model	
	Poor parent	Rich parent	Poor parent	Rich parent
No high-school	0.05	0.01	0.21	0.02
High-school	0.42	0.18	0.24	0.07
Some college	0.27	0.23	0.23	0.19
College degree	0.16	0.33	0.19	0.33
Graduate degree	0.10	0.25	0.13	0.40

Notes: Table entries are probabilities of obtaining a given schooling attainment (rows) conditional on parental endowment. Poor parents are those with log parental endwoment below the median. Rich parents are those with log parental endowment above the median.

Table 9: Occupational Choice Conditional on Parental Endowment and Schooling

Corr(data,model)	Poor parent	Rich parent
No high-school	0.68	0.27
High-school	0.85	0.76
Some college	0.64	0.14
College degree	0.64	0.57
Graduate degree	0.81	0.76

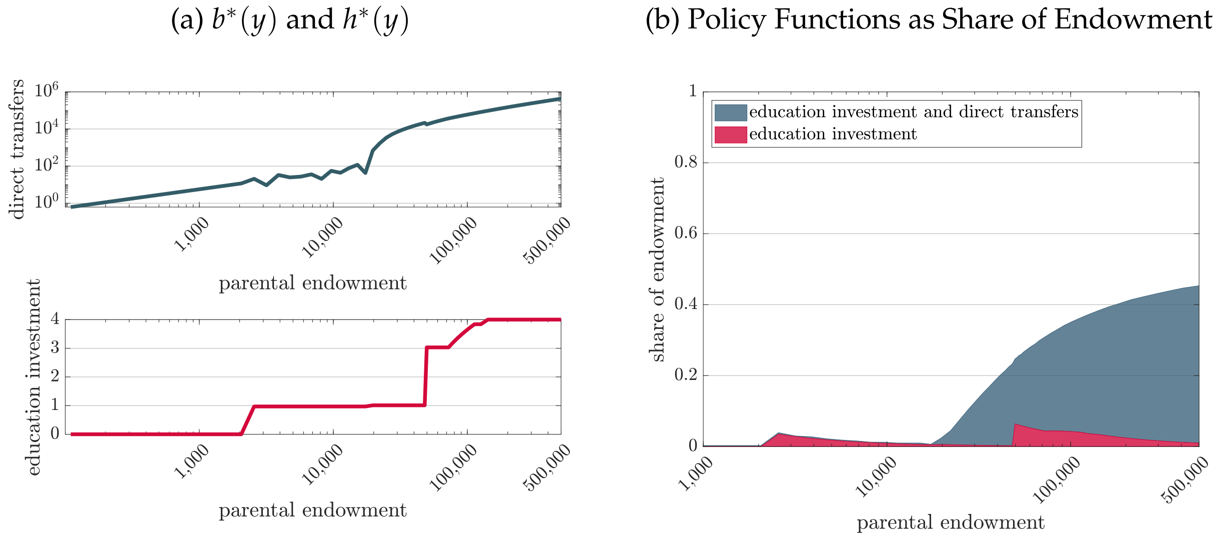
Notes: Table entries are correlation coefficients between occupational choice probabilities conditional on parental endowment and schooling predicted by the model and their counterpart in the PSID data. Poor parents are those with log parental endwoment below the median. Rich parents are those with log parental endowment above the median.

ble reports correlation coefficients between occupational choice probabilities conditional on parental endowment and schooling predicted by the model and their counterpart in the PSID data. These correlation coefficients are positive and high, suggesting that the model is able to capture what occupations are more likely to be chosen by children with a given educational attainment and parental endowment.

C.4.2 Policy Functions

Figure 16 displays the policy functions for education investment $h^*(y)$ and direct transfers $b^*(y)$. As Panel (a) of the figure shows, both direct transfers and education investment are increasing in parental endowment. Panel (b) shows that poor parents transfer resources to their children mainly by investing in their human capital. In contrast, rich parents devote a larger share of their endowment to direct transfers. We note that the apparent non-monotonicity in the policy function for the share of endowment spent on children's education simply reflects the discrete nature of our education groups. That this share is decreasing in parental endowment at high levels of parental endowment is a consequence of the fact that in the PSID data we only observe the number of years of schooling and cannot distinguish more refined aspects of schooling attainment such as the major or the quality of college education. In summary, the policy function under the estimated model broadly satisfies the main condition of the theory.

Figure 16: Investment in Education and Direct Transfers



Notes: Panel (a) shows the policy functions for direct transfers (top) and education investment (bottom). Panel (b) shows direct transfers and education investment as share of parental endowment.

C.5 The Decomposition of Persistence in Earnings

In this appendix, we provide a decomposition that characterizes the channels through which the model generates intergenerational persistence. Our model offers a simple characterization for the last measure in Table 3, i.e., the covariance between child earnings and parental endowment y. Let $\mathbb{C}_{ey}(\log e, \log y)$ denote the covariance between log earnings and log parental endowment:

$$\begin{aligned}\mathbb{C}_{ey}(\log e, \log y) &= \mathbb{E}_{ey}\left[\log e\ \left(\log y - \mathbb{E}_y[\log y]\right)\right],\\ &= \mathbb{E}_y\left[\mathbb{E}_e[\log e\,|y]\ \left(\log y - \mathbb{E}_y[\log y]\right)\right].\end{aligned}$$

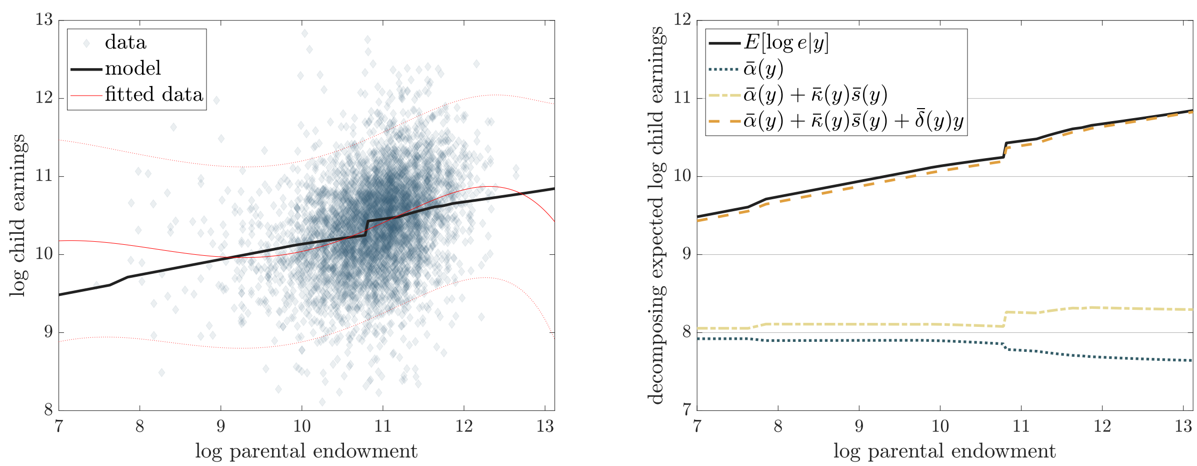
We can decompose the conditional expectation of the earnings of children given parental endowment to different components stemming from the dependence of the schooling and occupational choices of the former on the endowment of the latter. To build toward this decomposition, let us first define the conditional joint probability of occupational choice, talent, and schooling given parental endowment as

$$\mathbb{P}\left(j, s, u|y\right) \equiv \mu_j(s, u, y)\,\mathbb{P}_u(u)\,\mathbb{P}_{s|h}(s|h^*(y)), \tag{32}$$

where the conditional probabilities of occupational choice are given by Equation (9). Using this joint distribution, and with slight abuse of notation, we can define a number of

Figure 17: Child Earning vs. Parental Endowment

(a) Data vs. Model (b) Decomposition



Notes: Panel (a) compares the relationship between log earning and log parental endowment across child-parent pairs in the data. The red lines show a 3-degree polynomial fit and the corresponding 95% confidence bands. The solid black line shows $\mathbb{E}_e[\log e|y]$ implied by the model. Panel (b) decomposes the conditional expected log earnings of the children given parental endowment to different components based on Equation (33).

marginal conditional distributions. For instance, the conditional distribution of occupational choice given parental endowment is given by $\mathbb{P}(j|y) \equiv \int \sum_s \mathbb{P}(j,s,u|y)\,du$ and the conditional distribution of schooling given parental endowment is $\mathbb{P}(s|y) \equiv \int \sum_j \mathbb{P}(j,s,u|y)\,du = \mathbb{P}_{s|h}(s|h^*(y))$.

Based on the definitions above, Equation (14) implies that we can write the expected value of child earnings conditional on parental endowment as

$$\mathbb{E}_e[\log e|\,y] = \bar{\alpha}(y) + \bar{\kappa}(y)\,\bar{s}(y) + \bar{\delta}(y)\,\log y + \mathbb{C}_{js}\left(\kappa_j, s|\,y\right) + \mathbb{C}_{ju}\left(\theta_j, u|\,y\right), \tag{33}$$

where we have defined the expected values of the parameters of the earnings function conditional on parental income y, e.g., $\bar{\alpha}(y) \equiv \mathbb{E}_j\left[\alpha_j|y\right] \equiv \sum_j \alpha_j \mathbb{P}(j|y)$, and similarly for $\bar{\delta}(y)$ and $\bar{\kappa}(y)$. Similarly, we have defined the expected level of schooling conditional on parental income as $\bar{s}(y) \equiv \mathbb{E}_s\left[s|y\right] = \sum_s s\,\mathbb{P}_s(s|h^*(y))$, as well as the following two conditional covariances given parental endowment y:

$$\mathbb{C}_{js}\left(\kappa_j, s|\,y\right) \equiv \mathbb{E}_{j,s}\left[\kappa_j\left(s - \bar{s}(y)\right)|y\right], \tag{34}$$

$$\mathbb{C}_{ju}\left(\theta_j, u|\,y\right) \equiv \mathbb{E}_{j,u}\left[\theta_j u\,|\,y\right]. \tag{35}$$

The first term in Equation (33) captures the variations in the fixed component of earnings as a function of parental endowment, which captures the earnings of an agent with no

schooling ($s = 0$), a unit parental endowment ($y = 1$), and a mean level of talent ($u = 0$). As we saw in Section 4.2, the fixed component of earnings varies negatively with the returns to schooling and talent across occupations. The second term in Equation (33) accounts for the product of the conditional mean return to schooling and conditional mean schooling given parental endowment. This term captures two distinct forces: the patterns of occupational choice through which the children of rich parents may sort into occupations with higher returns to schooling, and the patterns of schooling attainment through which the children of rich parents receive higher educational investment and schooling. Similarly, the third term accounts for the mean return to parental endowment, capturing the potential sorting of the children of rich parents into occupations with higher returns to parental endowment.

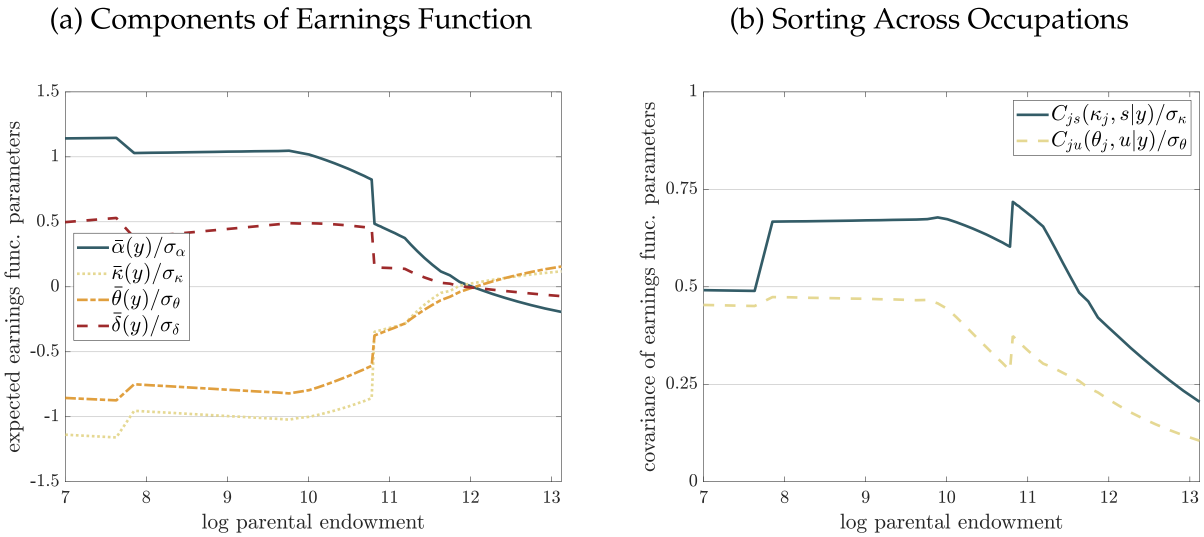
The last two terms in Equation (33) account for the patterns of sorting of children with higher schooling and talents toward occupations with higher returns to schooling and talent, respectively, *conditional* on parental endowment. The two covariances defined by Equations (34) and (35) capture how these two patterns of sorting vary across children with different levels of parental endowment. The stronger each of these two sorting patterns, the higher the conditional expected value of the log earnings of the children.

The Decomposition under the Benchmark Model Figure 17a compares the conditional expected earnings of children $\mathbb{E}_e[\log e|\, y]$ implied by the model with the observed relationship in our PSID sample. The model illustrates that the expected log earnings based on the model closely resembles that in the data. Accordingly, as reported in the last row of Table 3, the model comes very close to the observed covariance in the data. Figure 17b decomposes the expected log earnings in the model into different components following Equation (33). We find that the first three terms of the equation together explain the lion's share of the expected relationship between log earnings and parental endowment.

We find that the conditional expectation of fixed earnings $\overline{\alpha}(y)$ falls in parental endowment due to the fact that the children of richer parents sort into occupations with higher returns to schooling and talent and lower fixed earnings. Next, we find that the contribution of schooling $\overline{\kappa}(y)\overline{s}(y)$ increases in parental endowment, due to both the rise in the expected returns to schooling and the expected schooling attainment.[47] However, the estimation results suggest that through the lens of the model the main driver of the *variations* in expected log earnings as a function of parental endowment is the direct effect of parental endowment on earnings through the term $\overline{\delta}(y)\, y$. Despite sizable variations in the patterns of sorting

[47]Figure 18a in Appendix D shows how the conditional expected value of each component of the earnings function varies with parental endowment. We find that the expected returns to schooling $\overline{\kappa}(y)$ and to talent $\overline{\theta}(y)$ rise in parental endowment, while the returns to parental endowment $\overline{\delta}(y)$ fall in parental endowment.

Figure 18: Drivers of Persistence in Earnings



Notes: Panel (a) shows how the conditional expectation of different components of the earnings function across occupations vary with parental endowment. Each component is normalized by its corresponding standard deviation across the entire population, e.g., $\sigma_\alpha \equiv \mathbb{V}_j[\alpha_j]$ based on the stationary distribution of occupational choice. Panel (b) shows the normalized conditional covariances of schooling and returns to schooling, and talent and returns to talent.

across occupations conditional on parental endowment, Figure 17b shows that these variations make quantitatively negligible contributions to the overall dependence of expected log earnings on parental endowment.

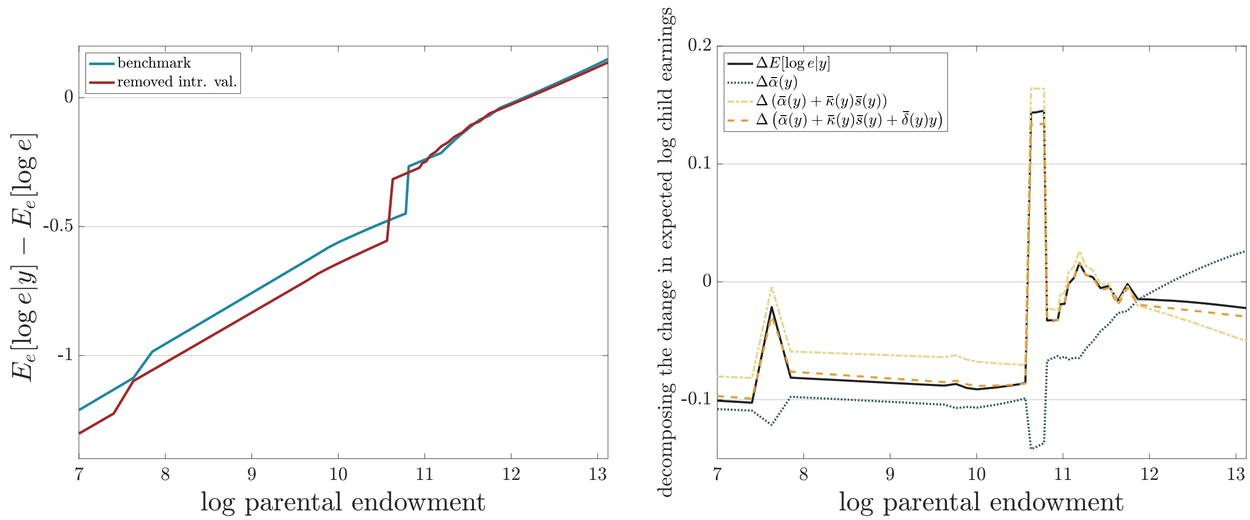
Figure 18b focuses on the two relevant patterns of sorting: the covariance of schooling and returns to schooling, and the covariance of talent and returns to talent. Both these two covariances are initially stable as parental endowment rises, but then eventually fall as parental endowment continues to rise. The reason is that the children of very rich parents become relatively more responsive to their idiosyncratic taste shocks and intrinsic quality of occupations and thus do not respond as strongly to the earnings incentives in their occupational choice.

Persistence of Earnings without Intrinsic Qualities As we saw in Table 4, the persistence in earnings slightly rises relative to the benchmark model when we remove the variations in the intrinsic qualities. Several forces together help shape this change in persistence. First, the most pronounced effect of removing intrinsic qualities for the children of the poorest and richest households is on the generate equilibrium response in the fixed component of their earnings. As we saw in Figure 8b, the wage rates fall in low-intrinsic quality occupations, chosen mostly by the children of the poorest parents under the benchmark, and rise in high-

Figure 19: Expected Log Earning vs. Parental Endowment, Removed Intrinsic Qualities

(a) Benchmark vs. Removed Intr. Val.

(b) Decomposition of the Change



Notes: Panel (a) compares the relationship between conditional expected log earning and log parental endowment between the benchmark model and that with removed variations in intrinsic qualities. Panel (b) decomposes the the change in the conditional expected log earnings of the children given parental endowment to different components based on Equation (33), in going from the benchmark model to the one with removed variations in intrinsic qualities.

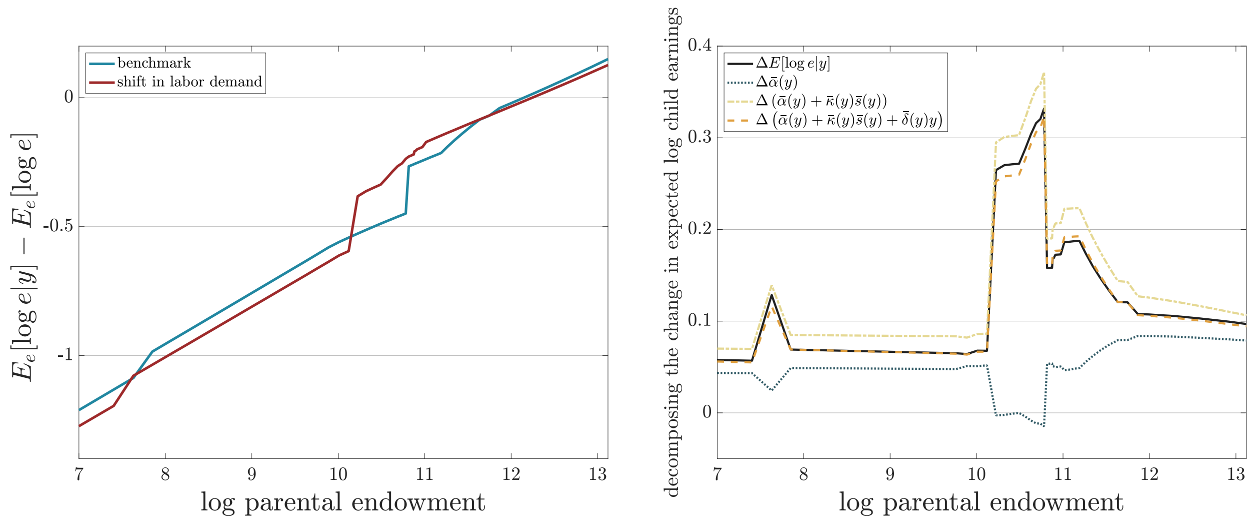
intrisic quality occupations, chosen by the children of the richest. To the extent that the children of poor children switch to occupations with higher intrinsic qualities, this further lowers the fixed component of their earnings due to the negative correlation between the intrinsic qualities and the fixed components of income α under the benchmark (see Table 2b). The most pronounced effect on the earnings of the children of middle class parents is through their schooling. These children are those most likely to switch to occupations with high intrinsic qualities, which happen to also have higher returns to schooling κ (see Table 2b). Their expected earnings rise due to higher schooling investment and attainment.

Figure 19a compares the conditional expected log earnings as a function of parental endowment under the benchmark with that under the model with removed variations in intrinsic qualities. Therein, Figure 19b decomposes the changes between the two conditional expectations to the different components based on Equation (33). We can see that the conditional expectation of the fixed component of log earnings $\overline{\alpha}(y)$ explains most of the differences between the children of the poorest and the richest parents, while the term involving the expected returns to schooling $\overline{\kappa}(y)\overline{s}(y)$ explains the change for the children of the middle class.

Figure 20: Expected Log Earning vs. Parental Endowment, Shift in Labor Demand

(a) Benchmark vs. Shifted Labor Demand

(b) Decomposition of the Change



Notes: Panel (a) compares the relationship between conditional expected log earning and log parental endowment between the benchmark model and that with shifts in occupational labor demand. Panel (b) decomposes the the change in the conditional expected log earnings of the children given parental endowment to different components based on Equation (33), in going from the benchmark model to the one with shifts to occupational labor demand.

Decomposition with Shifts in Labor Demand Figure 20a compares the conditional expected log earnings as a function of parental endowment under the benchmark with that under the model with shifts in labor demand. Figure 19b further decomposes the changes between the two conditional expectations to the different components based on Equation (33). The term involving the expected returns to schooling $\bar{\kappa}(y)\bar{s}(y)$ constitutes the main source of changes in expected log earnings.

D Additional Figures and Tables

Figure 21a compares the occupational choice elasticities estimated that control for potential earnings and the occupation of the parent against those without such controls, using the PSID data.

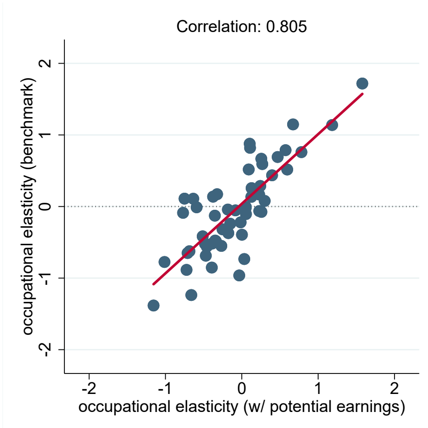
Figure 22 displays the correlation between endowment elasticities estimated with the PSID and NLSY data.

Figure 23 displays the relationship between occupational choice elasticities estimated with the NLSY data and the intrinsic quality of occupations.

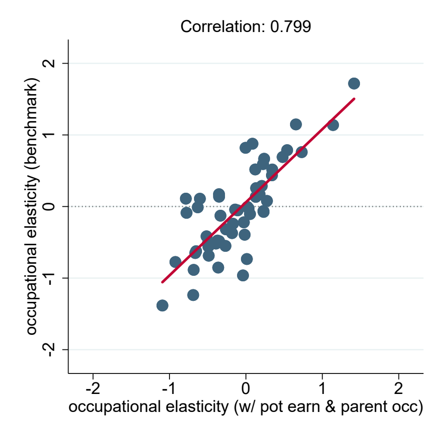
Figure 24 displays the relationship between occupational choice elasticities estimated

Figure 21: Occupational Choice Elasticities Controlling for Potential Earnings and Parent's Occupation

(a) Potential Earnings



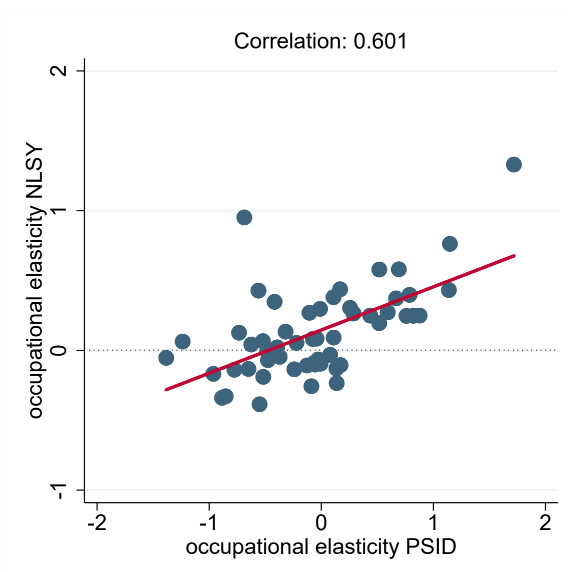
(b) Potential Earnings and Parent's Occ



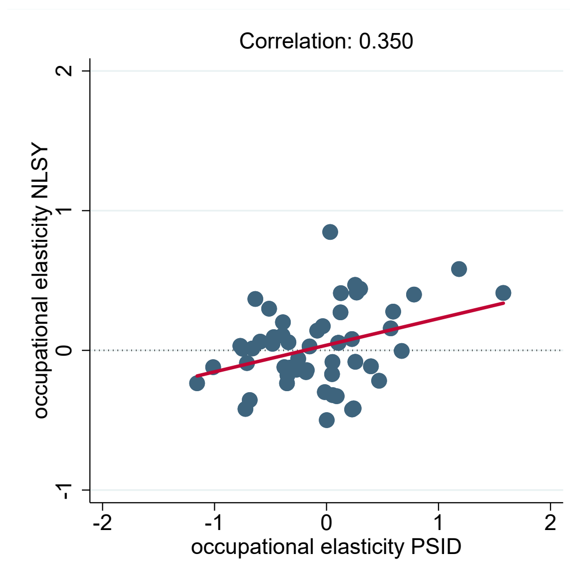
Notes: The left panel depicts the benchmark occupational choice elasticities against elasticities from the conditional logit estimation controlling for potential earnings across all occupations. The right panel depicts the benchmark occupational choice elasticities against elasticities from the conditional logit estimation controlling for potential earnings across all occupations and a dummy variable that is equal to one if the parent works in the given occupation.

Figure 22: Occupational Choice Elasticities, PSID vs NLSY

(a) Without Earnings Controls

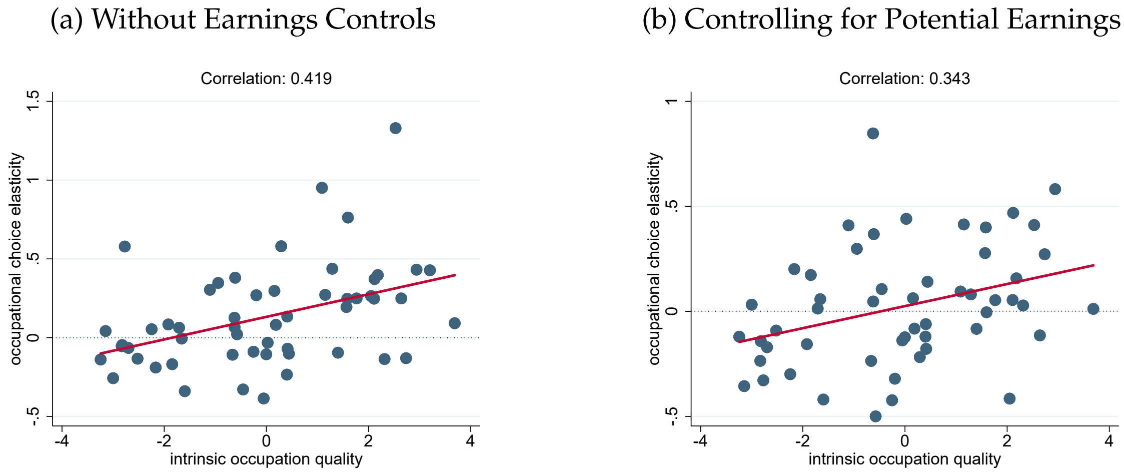


(b) Controlling for Potential Earnings



Notes: The left panel depicts the benchmark occupational choice elasticities. The right panel depicts the occupational choice elasticities estimated controlling for potential earnings in all occupations. The standard error of the correlation coefficient in the left panel is 0.111 and that of the correlation coefficient in the right panel is 0.131.

Figure 23: Occupational Choice Elasticities and the Intrinsic Quality of Occupations, NLSY



Notes: Panel (a) shows the relationship between occupational choice elasticities (vertical axis) and the intrinsic quality of occupations (horizontal axis). Panel (b) shows the relationship between occupational choice elasticities estimated controlling for potential earnings and the intrinsic quality of occupations. The standard error of the correlation coefficient in the left panel is 0.126 and that of the correlation coefficient in the right panel is 0.132.

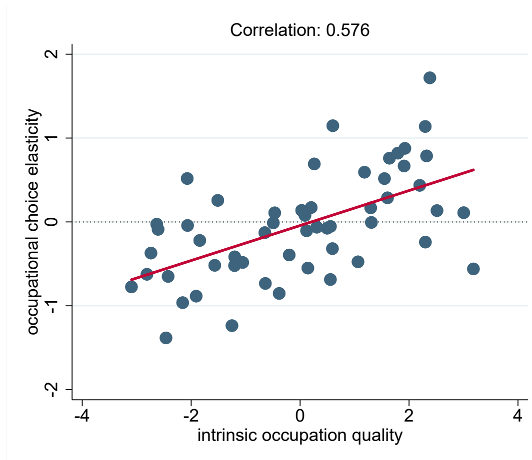
with the PSID data and the intrinsic quality of occupations under two alternative specifications. In the left panel, occupational choice elasticities and the intrinsic quality of occupations are estimated for the 80 occupation groups in Table 6. In the right panel, we maintain the occupation classification with 54 groups in Table 5, but define the intrinsic quality of occupations to be the first principal component of 5 job characteristics only: treated with respect, little hand movement, little heavy lifting, keep learning new things, do numerous different things. In both cases, the correlation remains positive, high (0.52 and 0.58, respectively) and statistically significant.

Table 10 examines whether controlling for the risk of occupations alters the relationship between occupational choice elasticities and intrinsic qualities. Column (1) reports results from projecting occupational choice elasticities on the intrinsic qualities of occupations. Columns (2) and (3) add to this projection a control for the coefficient of variation of log earnings by occupation, measured as the ratio between the standard deviation and the average log earnings by occupation. In column (2) the coefficient of variation of log earnings by occupation is calculated based on the pooled sample of the ASEC waves from 1976 to 2017. In column (3) the coefficient of variation of log earnings by occupation is calculated controlling for age (16-25, 26-35, 36-45, 46-55, 56-64), sex, race (white, Black, other) and year.

Figure 24: Occupational Choice Elasticities and The Intrinsic Quality of Occupations, Robustness

(a) Alternative Occupation Classification

(b) PCA 5 Job Characteristics



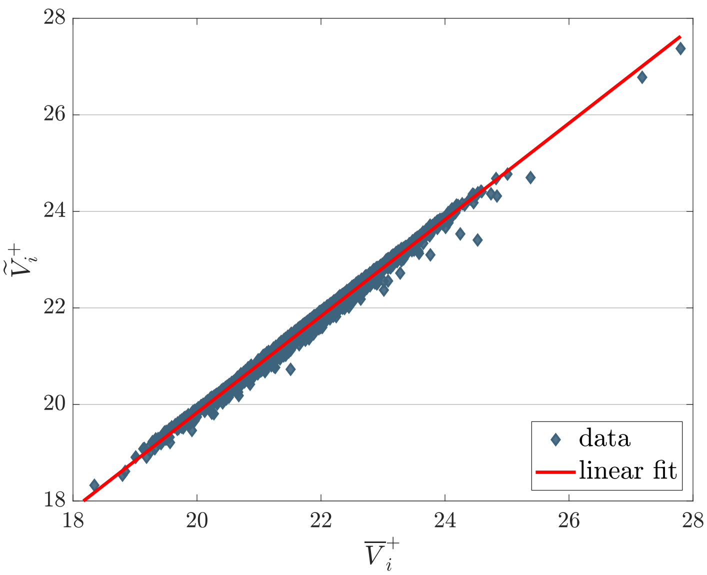
Notes: The left panel is based on an occupation classification with 80 occupation groups. In the right panel the intrinsic quality of occupations is estimating by applying the PCA on 5 job characteristics. The standard error of the correlation coefficient in the left panel is 0.097 and that of the correlation coefficient in the right panel is 0.113.

Table 10: Occupational Choice Elasticities, Risk and the Intrinsic Quality of Occupations

	(1)	(2)	(3)
Intrinsic quality, ν	0.197 (0.037)	0.183 (0.038)	0.188 (0.039)
Coeff. of variation log earnings		-3.101 (2.383)	-2.570 (3.396)
Constant	-0.043 (0.069)	0.183 (0.307)	0.319 (0.286)
Controls	–	No	Yes
R^2	0.352	0.372	0.359

Notes: The table shows the intercept, the slope coefficients and the R-squared of a regression of occupational choice elasticities on the intrisic quality of occupations (column 1) and the coefficient of variation of log earnings by occupation (columns 2 and 3).

Figure 25: Comparison Between the Two Welfare Measures $\overline{V}^{+}$ and $\widetilde{V}^{+}$

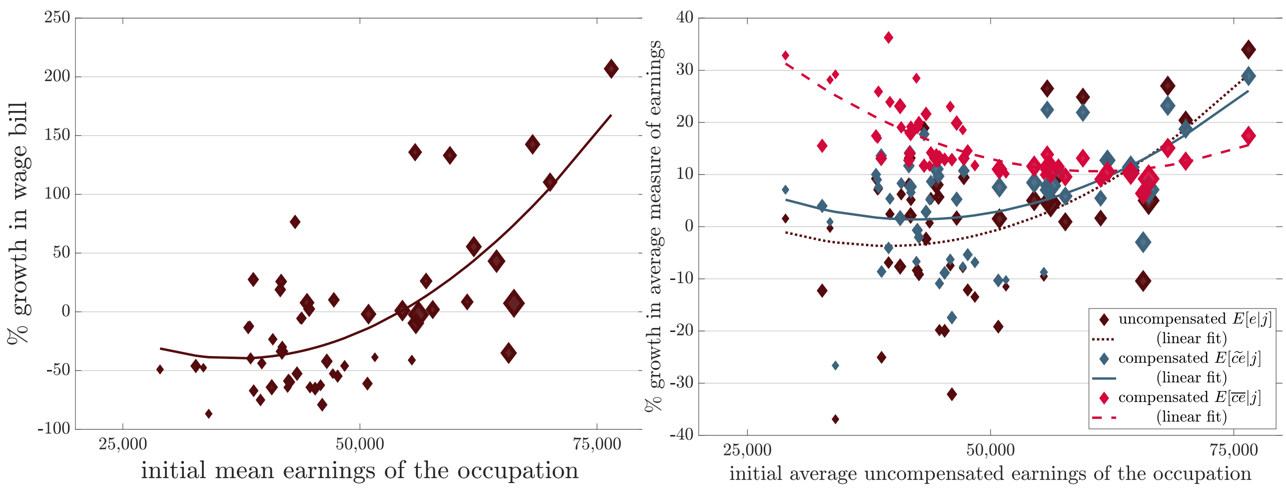


Notes: The figure displays a scatter plot of our two proxies for welfare of each child in our sample.

Figure 26: Growth Across Occupations

(a) Growth in Occupational Labor Demand

(b) Growth in Mean Earnings



Notes: Panel (a) shows the growth in occupational labor demand from the 1980–1985 average to the 2010–2015 average, as a function of the mean uncompensated earnings of the occupations under the benchmark model. Panel (b) plots the the growth in mean uncompensated and compensated earnings across occupations in response to the growth in occupational labor demand over the period, as a function of the mean uncompensated earnings of the occupations under the benchmark model.