Who Can Have It All?

Inter-generational Mobility & Unequal Cost of Career

Patrizia Massner Arash Nekoei *

March 20, 2025

Abstract

The standard measures of intergenerational mobility are based on occupation or income and thus overlook the costs of upward mobility. One such cost arises from the career-family tradeoff. Theoretically, greater costs (i.e., a steeper tradeoff) for individuals from disadvantaged backgrounds imply that income mobility overestimates welfare mobility. Empirically, Swedish administrative data shows that, indeed, to reach higher incomes, individuals with low-income parents sacrifice more of their family life. They delay having children, have fewer children, and spend less time with them. This disparity arises from labor market inequalities: to achieve high incomes, disadvantaged individuals require greater cognitive and non-cognitive skills, pursue longer studies in less lucrative fields, and thus start their careers later at lower-paying firms, while frequently changing jobs. We conclude by discussing how welfare mobility clarifies the link between the concepts of income mobility and equality of opportunity.

*Patrizia: IIES , patrizia.massner@iies.su.se , Arash: IIES and CEPR, arash.nekoei@iies.su.se. We thank Konrad Burchardi, Raj Chetty, Claudia Goldin, Markus Jäntti, Enrico Moretti, and Matt Notowidigdo for their comments.

1 Introduction

I heard somebody say one time: You can make it, if you try
And some of us, we tried so hard, we tried so hard.
Syl Johnson, 1969

Intergenerational mobility has emerged as a primary statistic of a society's fluidity and inclusivity in academic and policy debates.[1] A mobile society is deemed more equitable as it ensures equal opportunity for its members *and* more efficient since it discovers and harnesses talent regardless of background (Bell et al., 2019; Jones, 2022). [2]

The modern concept of intergenerational mobility first emerged in sociology (Sorokin, 1927). It describes the movement of individuals across generations between social classes, primarily measured using occupations. In economics, the focus has shifted to examining the correlation between the income or wealth of two consecutive generations.[3]

Yet, focusing on mobility in outcomes—such as occupation, income, or wealth—neglects the potential costs associated with upward mobility. Time and energy spent on career advancement and climbing the social ladder strain other aspects of life, such as family, leisure, and health, straining the delicate balance between personal and professional lives. A key element of this tension is the widely recognized career-family tradeoff (Goldin, 1997, 2021).

The ideal concept of mobility should extend beyond professional outcomes and capture overall welfare, thereby offering a more holistic perspective of mobility. However, a fundamental challenge is the lack of insight into individuals' preferences regarding their professional and personal lives.

This paper tackles this challenge by focusing on the costs of upward income mobility. We measure how much personal life an individual must sacrifice to cultivate a professional career and how this sacrifice varies by parental background. Thus, the paper extends the standard measures of intergenerational mobility by demonstrating that the career-family tradeoff is steeper for upstarts.

Consider an illustrative example. Despite coming from a disadvantaged, low-income background, Sofia became a first-generation college graduate. She joined a prestigious firm with up-or-out progression policy where she is a partner today. From the perspective of upward mobility, her achievements seem remarkable. However, compared to her peers, Sofia paid a significant personal price. The endless studying required in college to catch up strained her social life; she found a partner later in life and delayed starting a family. Therefore, from the standpoint of welfare mobility, her success is more complicated and ambivalent.

[1] See Harrington (1997); Solon (1999); Beaud et al. (2013); Chetty et al. (2014); Putnam (2016); Major and Machin (2018); Mogstad and Torsvik (2023) and references therein.

[2] See Corak (2006); Jäntti et al. (2006); Björklund et al. (2009); Blanden (2013); Chetty (2021).

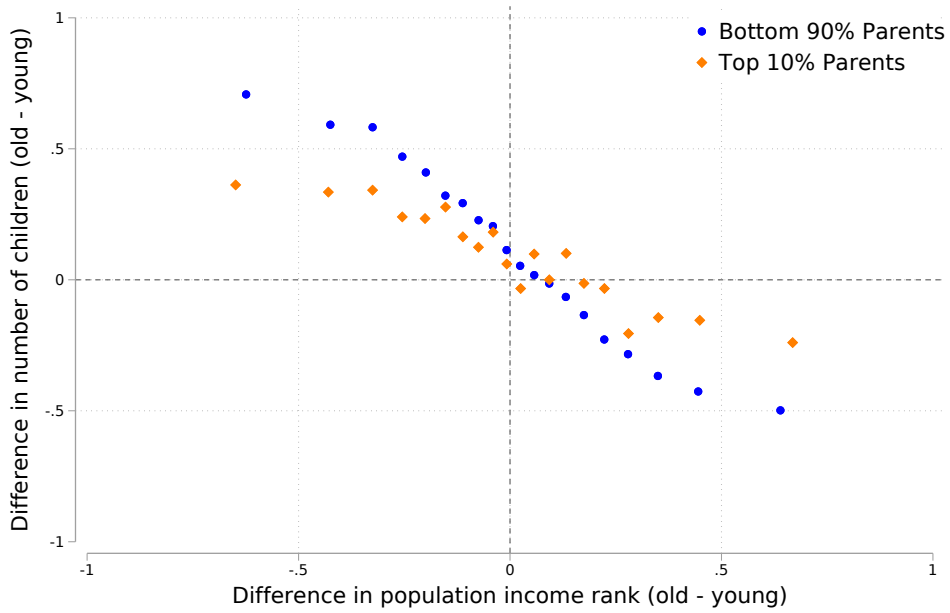
[3] See Nico (2021) for review of the literature in sociology and Björklund and Jäntti (1997); Solon (1999); Black and Devereux (2011); Mogstad and Torsvik (2023) for economics. Occupation is also the primary focus of historical studies of mobility, see Van Leeuwen and Maas (2010), Long/Ferrie 2013 Berger/et al. 2023.

We start with a theoretical model of intergenerational mobility. Agents receive utility from their professional and personal lives but face a career-family tradeoff dictated by resource (e.g., time or energy) constraints. Our main theoretical result is that income mobility overestimates welfare mobility if the cost of a career decreases by background. The idea is that upward mobility in professional life does not count as welfare mobility if it is achieved at the expense of the personal life.

Empirically, we test this pattern using Swedish register data. As a proxy for personal life, we focus primarily on family outcomes. Our main measures of family outcomes are the total number of children (completed fertility rate) and the time spent on parental leave per child. We complement this analysis with other aspects of personal life, such as health outcomes and geographical mobility. Professional life and parental background are measured through an individual's and their parents' lifetime income.

Figure 1 presents an illustrative example linking our theory to the data. Between two Swedish sisters, the one with a higher lifetime income tends, on average, to have fewer children. More importantly, this relationship varies by parental background. Among sisters whose parents are at the bottom 90% of the income distribution, moving from the bottom to the top of the income distribution results, on average, in having one fewer child. In contrast, for sisters with parents in the top decile of income, the difference is approximately half a child.

Figure 1: Income vs. Fertility of Sisters by Parental Background



Note: This figure compares sisters by plotting the difference in their lifetime number of children (completed fertility rate) against their lifetime income rank. The former is measured as all biological offspring born before age 45, and the latter as the sum of all labor income between ages 25 and 45, ranked within a birth cohort. We restrict the sample to 78,462 sister pairs. Top 10% (Bottom 90%) Parents indicate whether the parents belong to the top 10th decile (bottom nine deciles) of the sum of the father's and mother's income.

Building on this example, we empirically identify the cost of mobility by comparing the differ-

ence in family outcomes of individuals with and without careers by parental background. More precisely, we estimate the cross-derivative of family outcomes with respect to career and background. This approach is similar to a difference-in-differences design, where career serves as the treatment and background as the time variable, albeit without a causal interpretation. Formally, we regress family outcomes on two-sided fixed effects for an individual's own and parental income percentiles, along with the interaction terms of these two variables. These interactions capture the differential family costs of a career based on parental background.

The cost of a career decreases with parental background. To reach the top decile of lifetime income, those with low-income parents sacrifice more of their family life: They are two percentage points less likely to become parents. If they have children, they have them on average two years later and have .1 fewer children (corresponding to 5 percent fewer children). The extensive and intensive margins add up to .15 fewer children (7.5 percent). The difference in the intensive margin can be entirely explained by the postponement of parenthood. Controlling for the age of parenthood, self-made achievers have the same number of children as achievers coming from high-income families.

Self-made achievers not only have fewer children but also spend less time with each child. We demonstrate this by documenting a 10 percent reduction in the length of parental leave per child for female self-made achievers. Swedish parental leave is universally provided and funded by the social insurance system. However, collective bargaining agreements (CBAs) and firms often offer additional top-ups. To understand the role of these factors—along with occupations, age, and other variables—we investigate how much self-made achievers differ from privileged achievers along these dimensions.

The primary mechanism underlying the unequal family costs of mobility stems from the contrasting trajectories of upward mobility. Individuals born into affluent families often follow a streamlined and efficient path through education and career progression ("guided success"), which allows them to allocate more time to familial responsibilities. In contrast, self-made achievers typically endure numerous setbacks and recalibrate their paths repeatedly before achieving success ("hard-earned victory"), resulting in a more arduous and time-intensive journey. We show this with the help of our diff-in-diff approach and detailed data on education attainment, cognitive and noncognitive skills from the military, and labor market outcomes.

To reach the top income decile, individuals with low-income parents pursue an additional year of education compared to their counterparts with high-income parents. Furthermore, their military records show that they possess higher levels of both cognitive and noncognitive skills before entering university at age 18. This evidence suggests a lower return on education and skills for disadvantaged achievers. The lower return to years of schooling can be partially explained by the fact that they pursue less financially rewarding fields of study, as evidenced by the average income of students in these fields.

To compare the magnitude of these mechanisms, we convert the units of all four human capital measures to income rank. We regress income rank on education level and field as well as cognitive

and noncognitive skills. The predicted income ranks then provide a standardized measure for the different types of human capital.

Finishing their education later, disadvantaged achievers enter the labor market later in life. Moreover, they begin their careers at lower-paying firms, indicated by low AKM firm fixed effects (Abowd et al., 1999; Card et al., 2013). In an effort to catch up, they ascend the job ladder by frequently changing jobs and enhancing the quality of firms they work for, as evidenced by increasing AKM firm fixed effects. As direct evidence of impediment, self-made achievers are half as likely to work at the same company as their parents.

Working their way through the education system and labor market with more trial and error, as we have observed, self-made achievers tend to move further from their hometowns and parents compared to achievers from privileged backgrounds. This can be seen as another unequal cost of career advancement, which elevates the cost of parenthood.

It is

Our results build on a large literature on social (inter-generational) mobility.

In this debate, the Nordic countries are often lauded as paragons of mobility. Elhakeem et al. (2017)

The paper unfolds as follows. Section 2 delves into the cost of mobility from a theoretical perspective. Section 4 does the same from an empirical perspective after Section 3 shows the data and empirical design. Section 5 investigates the mechanisms behind the differences in family outcomes. In Section 6, we examine further cost measures of mobility beyond family outcomes measures. Section 7 discusses our findings. Section 8 studies the connection between welfare mobility and equality of opportunity before Section 9 concludes the paper. We relegate additional background material and auxiliary analyses to the Appendix.

2 Unequal Cost of Upward Mobility: Theoretical Framework

This section studies the theoretical link between income and welfare mobility. It conveys the core message as simply as possible, relegating the details and extensions to the Appendix A.

We consider an individual with preferences over personal life, h, and professional life, z. Both aspects of life use a L dimensional vector as input to their production functions: $z = \phi(x)$ and $h = \varphi(y)$. We assume both production functions are increasing and concave. Nature puts a limit on total input available $x + y \leq T$, where T is an exogenous L-dimensional vector. This constraint creates a career-family tradeoff. [4]

The individual lifetime optimization then can be written as:

$$\max_{x,y \in \mathbb{R}^L} U(z, h)$$

where $z = \phi(x)$ and $h = \varphi(y)$ and under the L constraints: $x + y \leq T$.

[4]This model nest the labor supply model, where input is one-dimensional time, $L = 1$, x is hours worked and y is leisure time, and production functions are $\phi(x) = wx$ and $\varphi(y) = y$.

Similarly to converting firms' profit maximization into cost and scale optimization in production theory, we reformulate the individual decision-making problem into two steps: first, balancing the trade-off between professional and personal lives (family vs. career); and second, optimizing input vectors required for each aspect of life. Namely,

$$\max_{z} U\left(z, H\left(z\right)\right)$$

where the function $H\left(z\right)$ is the optimal personal life given a career level, which is defined as

$$H\left(z\right) = \max_{y} \varphi\left(y\right)$$

under professional life production function $\phi\left(x\right) = z$ and the constraint $x + y \leq T$. Note that $\frac{\partial H}{\partial z} < 0$, capturing the loss of personal life for one unit of career gain.

The role of parental background can manifest in three ways: it may shape the professional or personal life production functions, or it may alter the relevant constraint, expressed as $\phi(x, z_{-1})$, $\varphi(y, z_{-1})$, or $T(z_{-1})$. These three channels determine how the function H depends on z_{-1}. The key assumption is that parental career outcome serves as a sufficient statistic for the parental impact on individual welfare.

Thus, individual career vs. family optimization becomes:

$$\max_{z} U\left(z, H\left(z, z_{-1}\right)\right)$$

which is biased in background if $\frac{\partial^2 H}{\partial z \partial z_{-1}} < 0$.

Our main goal is to investigate how a biased background affects the career-family tradeoff and create a wedge between income and welfare mobility. We assume a steady-state economy. More precisely, all generations are identical with stable marginal and joint distributions across generations.

We address our research question using two complementary approaches. First, we provide an analytical analysis under specific functional forms. Second, we adopt an approximation that relies on certain distributional assumptions about z. In particular, we assume a symmetric distribution yielding a zero third central moment and a joint normal structure, allowing us to employ Isserlis' theorem.

In our first approach, We assume Stone-Geary preferences of over two dimensions of life, that is,

$$U(z, h) = \alpha \ln\left(z - \gamma_z\right) + \overline{\alpha} \ln\left(h - \gamma_h\right)$$

We further assume the optimal personal life function as

$$H\left(z, z_{-1}\right) = \left(1 - \frac{z}{w}\right) z_{-1}^{\beta}$$

where productivity follows a Markov chain where (w, w_{-1}) is jointly log-normal distributed with

correlation ρ.[5] If we assume that $\gamma_z = \gamma_h = 0$, then we can analytically address our question.

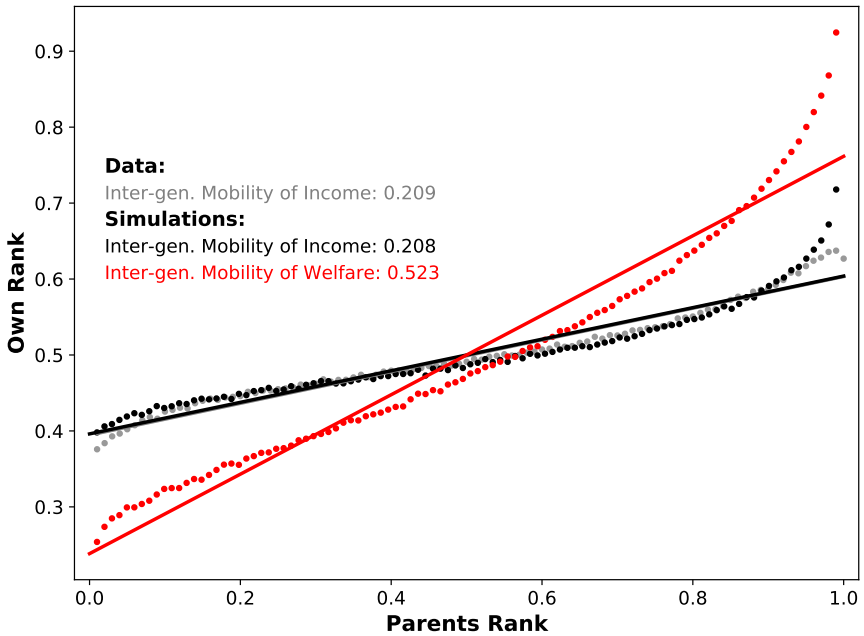
Proposition 1. *In the case where $\gamma_z = \gamma_h = 0$, the intergenerational welfare mobility is lower than the intergenerational income mobility when there is a biased background $\beta > 0$.*

The intuition is as follows. Given the absence of the income effect, the intergenerational mobility in income rank is $\frac{6}{\pi} \arcsin\left(\frac{\rho}{2}\right)$, because $z = \alpha w$. But the intergenerational mobility in welfare is different as the welfare rank depends on the ranking of a function of own and parent productivity, more precisely

$$\alpha \ln(w) + \overline{\alpha}\beta \ln\left(w_{-1}\right)$$

For the case where $\gamma_z, \gamma_h > 0$, we use a simulation. The results are depicted in Figure 2 below and the details are relegated to the Appendix. The parameter ρ is set so that income Spearman rank correlation matches the data.

Figure 2: Model Prediction: Intergenerational mobility overestimates welfare mobility



Note: This figure plots the rank-rank correlation of income (in black) and welfare (in red) in simulated data with .... We further plot the rank-rank correlation of income using Swedish register data described in Section 3. The slide lines are ... The number of observations are the same in the real and simuated data.

In our second approach, we use Taylor expansion to approximate th utility function U and the summary production function H.

$$H\left(z, z_{-1}\right) \simeq \alpha z + \beta z_{-1} - \gamma z_{-1} z$$

Symmetric distribution to have zero third central moments and joint normal using Isserlis' theorem. In this way, we are assuming that functions U and H are so that Individual optimizes,

[5]The case of $\beta = 0$ coincides with labor supply model at wage rate w.

leading to the distribution of z and a Markov chain which is stationary.

The intergenerational welfare mobility is lower than the intergenerational income mobility when we have a biased background $\frac{\partial^2 H}{\partial z \partial z_{-1}} < 0$.

3 From Theory to Data: Empirical Design and Data

3.1 Empirical Design

We empirically measure the unequal cost of career, defined in the theoretical section as $\frac{\partial^2 \widetilde{K}}{\partial c \partial c_{-1}} < 0$, by regressing lifetime family outcome of individual i, K_i, on individual's lifetime income, C_i, parental lifetime income, B_i, and the interaction of these two elements. More specifically, we run the following regression where we flexibly control for individual and parental lifetime income using gender-specific dummies of rank percentiles and linear trends:

$$k = f(c) + g(c_{-1}) + \sum_{i=1}^{9} \beta_i \underbrace{R(c)}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(D(c_{-1}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{1}$$

where $D(.)$ is decile function and $R(.)$ the rank function. Our coefficients of interest are $\{\beta_l\}$, capturing the interaction terms between individual and parental income. $\mathbb{1}(R_{10}(C_i) = 10)$ indicates having a lifetime income in the top decile, which we refer to as "having a career". We interact this term with parents' lifetime income deciles, $\mathbb{1}(R_{10}(B_i) = l)$, where we omit the highest decile as the reference group. Then, the coefficients of interest $\{\beta_l\}$ measure the difference in family outcomes between those with a career from low-income parents vs. those with a career from top-decile income parents minus the same difference for those without a career. In this way, our design captures a difference-in-differences parameter, albeit without any causal interpretation. A negative coefficient implies that individuals from a low background sacrifice more family life to achieve a career compared to individuals coming from a high background. Hence, $\{\beta_l\}$ measures the difference in the cost of upward mobility by background.

Alternatively, we comparing bottom 90% with top 10% parents:

$$k = f(c) + g(c_{-1}) + \beta R(c) \times \mathbb{1}(R_{10}(c_{-1}) \neq 10) + \varepsilon \tag{2}$$

We will also use a sibling comparison closer to Figure 1 using Parent FEs instead of $g(c_{-1})$

$$k = f(c) + \delta_p + \beta R(c) \times \mathbb{1}(R_{10}(c_{-1}) \neq 10) + \varepsilon \tag{3}$$

3.2 Data and Estimation Sample

We use individual-level data on the universe of the Swedish population from several administrative registers on (i) labor, capital, and business income spanning from 1968 to 2020 from the Income tax database, (ii) family relations from the Multi-generational register, (iii) demographic

information from The longitudinal integrated database for health insurance and labour market studies *LISA* and the register of the total population *RTB* spanning from 1990 to 2020 for the former and 1968 to 2020 for the latter, (iv) cognitive and non-cognitive skills from Swedish draft data, (v) firm characteristics from Employer-employee matched dataset, (vi) wages and hours from the Wage Structure Survey, and (vii) test scores at high school from the Educational register.

We construct several variables, which we define in the following. Individual lifetime income rank, C_i, is measured as the sum of labor income over the ages 25 to 45. We rank the income within the birth cohort of the individual. If individuals have the same income level, we assign the average rank of the group following Chetty et al. (2014).

Parental income rank, B_i, is measured as the sum of the total income of the mother and the father. Total income contains labor income but also business and capital income. We sum parental income over the individual's childhood, i.e., the ages 0 to 20.[6] Then, we rank incomes within the birth cohort of the individual following Chetty et al. (2014). We assign the average rank of the group in case of the same income levels.

We measure personal life outcomes through family-related indicators. Our primary focus is on the number of children in life. This is quantified as the number of biological children born before age 45, which serves as a proxy for the Completed Fertility Rate. We refer to it as a proxy because the analysis is restricted to age 45 to incorporate a broader range of cohorts into the study.[7] Unlike the commonly used Total Fertility Rate, which reflects the average number of children a woman would have if she experienced the current age-specific fertility rates throughout her reproductive years, the Completed Fertility Rate directly measures the total number of children born to an individual by the end of her reproductive span.

We complement this measure of a number of children with further family-related outcomes, such as the age at the birth of the first child and the average days of parental leave per child.

We restrict the analysis to individuals born in Sweden between 1955 and 1975. This restriction enables us to construct a balanced sample of individuals reaching at least the age of 25 by 1968 (the first year of income data) and 45 by 2020 (the last year of the data). We exclude individuals who die before the age of 45, have non-identifiable biological parents, or have parents born outside of Sweden. Furthermore, we drop individuals who receive any disability insurance or reside outside Sweden between the ages of 35 and 45.[8] This data restriction results in 1,590,709 individuals. In Figure A.1, we show that the selected sample has a similar time trend to the entire population of Sweden, suggesting that the sample restrictions do not lead to selection. In Table A.1, we report sample characteristics of our main variables. On average, individuals have 1.9 children before age 45, are 29 years old at the first birth, and take, on average, 220 days of parental leave per child.

[6] If individuals are born outside the period for which we have data, we use a shorter period. For example, for the cohort born in 1967, we use ages 1 to 20; for the cohort born in 1966, we use the ages 2 to 20, ..., up to the cohort born in 1955, for which we use ages 13 to 20.

[7] Appendix Figure A.1 provides further details on the cohort selection.

[8] We start at age 35 due to data restriction of disability insurance and residence status, which are variables of the LISA dataset beginning in 1990.

4 Unequal Cost of Upward Mobility: Empirical evidence

Figure 3 displays the costs of mobility measured by life-time number of children (Completed Fertility Rate) as our main outcome variable. Figure 3a focuses on the total number of children, while Figures 3b and 3c separate the results from extensive and intensive margins, respectively.

Reaching the top decile of income is associated with .15 fewer children for those from underprivileged backgrounds (the bottom 90% deciles of parent income) relative to those from affluent backgrounds (whose parents are from the top decile). The patterns are similar for both the extensive and intensive margin, with the magnitude of the effects indicating a nearly equivalent contribution from each margin.[9] The impact is slightly less pronounced for men than for women, largely attributable to differences in the extensive margin.

Underprivileged achievers tend to embark on family life at a significantly later stage. Figure 4a elucidates this phenomenon within the same analytical framework as previously employed. In this instance, the dependent variable is the age at the time of the first child's birth. Women from underprivileged backgrounds who ascend to the top decile of income are, on average, more than two years older at the birth of their first child compared to their counterparts who achieve similar income levels but hail from affluent backgrounds. For men, although the effect remains noteworthy, it is less pronounced. Looking into the exact timing of childbirth, we find that the difference stems from underprivileged achievers having their first child much less frequently before age 25 but instead much more often after age 35 (See Appendix Figure A.4).

The delay in childbearing accounts for the entire disparity in the number of children at the intensive margin. Figure 4c illustrates that, after adjusting for the age at first childbirth, the fertility difference between privileged and underprivileged achievers vanishes.

Achievers from underprivileged backgrounds not only tend to have fewer children on average but also dedicate less time to their children. Figure 4b reveals a significant difference in the amount of parental leave per child. Selfmade achievers take approximately one month less in parental leave per child than those from privileged backgrounds. This observation is even more significant because parental leave is universally available in Sweden, underscoring the contrast in leave utilization among different socioeconomic groups.

More precisely, Swedish parental leave is universally provided and funded by the social insurance system. However, collective bargaining agreements (CBAs) and firms often offer additional top-ups.

5 Mechanisms: Hard-Earned Victory vs. Guided Success

Two hypotheses may elucidate why underprivileged achievers tend to commence family life later and dedicate less time to it: first, the educational pathway is more uneven and time-intensive for

[9] The magnitude of the percentage change of the intensive margin is similar to the extensive margin: $4\% \times 2.1 \approx .1 \times .86$.

those from disadvantaged backgrounds; second, they encounter greater challenges in job searching and developing their professional careers, which consume significant amounts of their time and energy. This section aims to explore these hypotheses in detail.

5.1 Longer Education, Better Human Capital

The initial evidence is simple yet powerful: to reach the top decile, the underprivileged, on average, require one additional year of education compared to approximately 15 years for privileged achievers. Figure 5a illustrates this using the framework outlined in Section 3.1 (Regression 1) and examines years of education as the outcome.

Another critical aspect of education, often obscured due to data limitations, is the field of study. Utilizing a rich Swedish dataset on education level and field, we decompose the differences in the quantity and quality of education. Regarding the level, we exploit 47 distinct categories of highest completed education, ranging from pre-school level to having a doctoral degree (Appendix Table A.4). Regarding the field of study, we observe 336 categories, e.g., "Teacher training, industry and crafts" or "Industrial and product design" (Appendix Table A.3).

Moreover, we also use the skill level from Swedish draft data. For men, cognitive and non-cognitive skills are measured at age 18 as part of the military draft (Edin et al., 2022; Almgren et al., 2023; Hermo et al., 2022). We use four measures of cognitive skills and four measures of non-cognitive skills. Cognitive skill measures are based on test scores for (i) inductive skills, (ii) verbal comprehension, (iii) spatial ability, and (iv) technical understanding. Non-cognitive skill measures stem from an interview with a certified psychologist. They include scores for (i) emotional stability, (ii) intensity, (iii) psychological energy, and (iv) social maturity. These skill measures are only available for those who participated in military tests, i.e., men. Hence, we perform the regression only for men. We perform the equivalent exercise when omitting skill measures and report the respective results for men and women in Appendix Figure A.5.

Using these four measures of human capital—education length (level), field of study, cognitive skills, and non-cognitive skills—we can address questions such as: How do self-made achievers differ from their privileged peers in their field of study? What about their differences in cognitive or non-cognitive skills?

The challenge of using four measures of human capital lies in their different measurement units. To address this, we standardize them by converting them into income-equivalent units. This involves regressing an individual's lifetime income rank on these measured facets of human capital using a non-parametric approach with fixed effects for each measure. To avoid over-fitting, we run the above regression in the population, not the estimation sample. The regression has an R-squared of 0.247 over 558,109 observations. The predicted income ranks from the different measures of human capital serve as respective outcome variables of our main specification in regression 1. Figure 5b illustrates these findings.

To reach the top decile, self-made individuals exhibit higher skills, encompassing both cognitive and non-cognitive aspects, by age 18. They opt for fields of study with lower average

earnings but engage in longer periods of education. The beneficial impact of their extended education is roughly equivalent to the diminished earnings associated with their chosen fields. Thus, self-made individuals offset the lower financial returns of their fields by pursuing lengthier educational paths. Consequently, when accounting solely for educational outcomes, all achievers are predicted to have similar incomes. The distinguishing factor, then, lies in their skill levels. The superior cognitive and non-cognitive skills of self-made individuals compensate for either the unobserved skills of their more affluent peers or the advantages those peers have, such as social networks and parental know-how. These aspects outside of an individual's measured human capital are captured in the residual component.

5.2 Steep Job Ladder

Taking more time for their education, how do underprivileged achievers catch up on the labor market? In Figure 6, we report different dimensions of their working life. First, Figure 6a documents that selfmades start their career at later ages. To reach the top decile of the income distribution, those with parental income in lower deciles are, on average, half a year older at the time they start their first job than their high-income peers from advantaged backgrounds.

Furthermore, selfmades start their career in lower-paying firms judged by AKM firm fixed effects.[10] The first job of high-earners with lowest-decile parents has an earnings premium of four percentage points lower than high earners with highest-decile parents (Figure 6b). However, selfmades change jobs more often (Figure 6c), and their firm fixed effects grow more than their peers and with such a magnitude that by age 45 selfmades catch up with their peers (Figure 6b). This is because the magnitude of the difference in their first firm is equal to the magnitude of the growth.

This dynamic is also reflected in wages. Figure 6d displays the difference in wages for underprivileged achievers in the beginning vs. the end of their working life before age 45. While they have, on average, higher wages throughout this period, wages are especially higher later in their career with a six percentage difference to their privileged peers. Interpreting wage as a measure of productivity, this result reinforces the human capital results of Section 5.1 suggesting that individuals who climb the job ladder from low backgrounds are relatively more able and productive.

As we observe higher wages at equal earnings, we can expect differences in hours worked between the two groups. In Figure A.6, we report for the subsample of workers in the Wage Structure Survey actual, contracted, and overtime hours. All three measures of hours show the same pattern: Selfmade achievers work on average 3 percent less than their privileged counterparts.

Self-made achievers tend to possess higher levels of human capital, suggesting that privileged achievers may rely on additional mechanisms to enhance their career. One such mechanism is

[10] We compute the AKM firm fixed effect by regressing log earnings on individual fixed effects and firm fixed effects including controls of age and education following Abowd et al. (1999); Card et al. (2013). We run this regression for the working population of Sweden using the largest connected set of firms. We interpret the firm fixed effect as the average firm pay premium.

access to better social networks, which can assist them in identifying high-paying educational opportunities and firms.

To investigate this hypothesis, we construct a proxy for family networks. Following the employer inheritance literature (see Corak and Piraino (2011)), Figure 8a illustrates that self-made achievers are six percentage points less likely to work in their father's firm—a substantial effect, representing nearly 50 percent of the baseline likelihood. There is only limited evidence of reliance on classmate networks, which selfmades use to compensate for the lack of family networks (Appendix Figure A.3). This pattern supports the notion that privileged achievers can leverage additional resources beyond conventional human capital.

6 Private Life: Beyond Family

The employment trajectories of underprivileged achievers reveal a tendency to switch jobs more frequently before establishing their career paths, a trend that is also evident in their geographic mobility. Figure 7 explores these patterns.

Figure 7c shows that the average number of total moves for underprivileged achievers is 0.08 higher than on average 0.8 moves for their privileged counterparts. This goes along with moves that separate selfmade individuals from their families. We measure this two-fold: First, Figure 7a indicates they are 0.2 percentage points more likely to leave their birth county. Second, they live further apart from their parents later in life. In particular, at age 35 individuals who have a career and come from low-income parents live on average 55 km further away from their mother's residence compared to individuals from high-income parents. Finally, we investigate from which municipalities they are moving. Figure 7a reveals that selfmade achievers are born into higher-ranked municipalities. One interpretation is that in order to achieve the same career as their counterparts from high-income parents selfmade achievers need a high-income environment. One specific aspect of this argument—networks—are discussed in the prior Section (see Figure 8a).

This suggests that individuals from less advantaged backgrounds often need to exhibit greater geographic flexibility, sometimes at the cost of leaving their hometowns to secure career advancement. Nakamura et al. (2022) suggest a significant disutility of geographical moves. This is another aspect of personal life affecting welfare mobility.

We use family outcomes as our primary measure of private life. Yet, there are other aspects of life not captured in this paper due to data limitations. This includes, for example, time spent in leisure or investments in health. We address the latter by leveraging Swedish register data on healthcare utilization. We measure health outcomes along two dimensions. First, we show in Figure A.9 and Table A.2 that we do not find a statistically significant correlation on mortality. Second, we estimate Equation 1 using the total days spent on sick leave as an outcome variable. Selfmade achievers have on average 1.5 fewer days of sick leave compared to their privileged counterparts (Figure A.10a). To address the concern that we rather capture differences in healthcare utilization than actual differences in health between these two groups, we further report the within-person

difference in sick leave days over the working life. Doing it this way, we do not find a statistical difference between the groups (Figure A.10b). While this evidence suggests that health outcomes of selfmade achievers are not worse than privileged achievers, we have to note that estimates are not precisely estimated meaning that we cannot conclusively rule out any effect on health.

7 Discussions

7.1 Preference Heterogeneity and Within-family Analysis

Economists often avoid delving into the complexities of taste heterogeneity, a caution rooted in sound reasoning, as initially discussed by Gary Becker. Within our context, if taste heterogeneity exists, individuals who climb the social ladder are distinguished by their high productivity or low preference for personal life. Intuitively, given the presence of obstacles to success for individuals from disadvantaged backgrounds, those who manage to overcome these barriers and ascend likely possess not only higher productivity but also a lower propensity for personal life compared to their professional lives. This suggests that such achievers could either be exceptionally talented or inclined towards workaholism.

We address this issue by adding family fixed effects to our main regression in Equation 1. The idea is that if preferences for personal life are correlated within a family, we cancel out the average preference for the number children between siblings. Specifically, we add to Equation 1 father fixed effects to measure differences in family outcomes between siblings instead of the controls of parental income. We run complementary regressions using the biological mother to identify siblings within one family. By adding father (or mother) fixed effects, we control for not only parental lifetime income and within family preferences for the personal life but also further unobserved factors that are constant between siblings, such as the business network of parents or parental educational investments into the children. Furthermore, we include controls X_i in a subset of regressions. These controls contain county of birth, year of birth, or the number of siblings.

Figure **??** shows the results where we add father fixed effects instead of parental income ranks in Equation 1.[11] Our main result on the number of children is robust to the inclusion of family fixed effect, despite the magnitude being slightly lower. On average, individuals who have a career have 0.1 fewer children when coming from low-income parents. This suggests that preferences over private and working life cannot explain the entire differences in the number of children between selfmade and privileged achievers.

7.2 Quantity vs. Quality of Children

We saw that underprivileged achievers have, on average, fewer children. On the one hand, this could imply that they can spend more time per child, balancing the "quantity vs quality" tradeoff.

[11] For results by gender of individual or parents, see Appendix Figures **??** and **??**.

This would be in contrast to evidence we saw on parental leave (Figure 4b). On the other hand, if underprivileged achievers spend more time building their careers, this could crowd out family time, affecting both the number of children and the time spent per child. Data sources to measure time spent with the family are limited. Hence, we proxy human capital investments into their children using achieved test scores in school. Specifically, we report the probability of the offspring passing all three subjects (English, Swedish, and Maths) in the national test exam in grade 9.

Figure A.11a shows that the first child of underprivileged achievers is six percentage points more likely to pass all three subjects in the national exam compared to children from privileged achievers. This could either mean that underprivileged achievers invest more into the human capital of their (fewer) children (the channel we aim to test for) or that these children are more able.

To address the latter channel, we control in Figures A.12 and A.13 for parental skills assuming that there is a correlation between parental and children's skills. The magnitude of the coefficient drops to two percentage points when controlling for parents' education level and field and to one percentage point when further controlling for the father's cognitive and non-cognitive skills. In the latter case, most coefficients turn statistically insignificant. This evidence points towards the positive correlation being driven by children of selfmade achievers being more on average able. Conditional on parental skills, underprivileged achievers do not significantly invest more or less into their offspring.

7.3 Marriage & Assortative Mating

If individuals are facing a tradeoff between time spent with their family and having a career, the career of the spouse matters, especially as raising children can be done in a household. A.8 shows that underprivileged achievers mate with spouses that have, on average, higher ranks compared to achievers from high backgrounds. This pattern goes along with the results of high achievers spending less family time, as the extra time spent on building their career is not offset by the spouse spending more time at home.

7.4 Heterogeneity by Gender

The cost of mobility, measured as differences in family outcomes, is similar across genders. To begin, Figure **??** presents the main result of Regression 1, with the outcome being completed fertility. This is analogous to Figure A.2, but here the interaction term varies by gender. Self-made women have, on average, slightly fewer children than self-made men, but the difference is economically small and statistically insignificant. This difference is primarily driven by variations in the extensive margin of having children.

This pattern extends to the reported mechanisms. For instance, Figure A.5 illustrates the human capital trajectories of self-made achievers relative to privileged achievers, disaggregated by gender. Both self-made men and women pursue longer educational paths in less lucrative fields

of study. However, a notable gender difference emerges in parental leave days per child or the age at birth, as depicted in Figure 4 and discussed in Section 4.

7.5 Time Trend

We examine the change in the cost of mobility across 21 birth cohorts by dividing our sample into two periods: individuals born before 1965 and those born afterward (Figure A.7). The number of children among underprivileged achievers is negative for both cohort groups; however, the severity decreases over time. Specifically, underprivileged achievers born between 1955 and 1965 have, on average, 0.15 fewer children, compared to 0.1 fewer children for those born between 1965 and 1975. While the difference is not statistically significant for the bottom decile, the change in magnitude may reflect shifts over this time period.

8 Intergenrational Mobility vs. Equality of Opportunity

The evidence provided so far, both theoretical and empirical, reveals that background-biased mobility costs imply that intergenerational mobility (IGM) of output overestimates the IGM of welfare. This section contrasts these IGM notions with equality of opportunity (EOp) to elucidate their interrelations. It is not our purpose to debate the superiority of any particular concept from a normative standpoint. Instead, our aim is descriptive, focused on understanding these concepts and exploring their interconnections.

8.1 Definitions of Concepts

We start with a generalized version of our model in Section 2. Individuals decide a vector x, which we refer to as effort, according to

$$V(\sigma) = \max_{x} U\left(Y(x,\sigma), K(x,\sigma)\right) \tag{4}$$

given a vector of exogenous parameters $\sigma = \{\sigma_i\}_{i\in I}$.

Equality of outcome is achieved when all individuals have the same outcome, $\frac{dY}{d\sigma_i} = 0$ for all $i \in I$. In contrast, equality of welfare requires $\frac{\partial V}{\partial \sigma_i} = 0$ for all $i \in I$. Note that we focus on the idealized forms of each concept while abstracting away from the metrics utilized to gauge deviations. This approach adds a level of simplicity that aids in elucidating the connections between concepts. Subsequently, we discuss some generalizations in certain cases.

We will partition the space of exogenous parameters into four groups. First, to talk about IGM, we differentiate them based on whether they are hereditary or not, $I = \Sigma^1 \cup \Sigma^0$. We denote the implied sub-vectors as σ^1 and σ^0, so by definition, the hereditary vector is the same across generations, $\sigma^1 = \widetilde{\sigma}^1$. This distinction is similar to our framework of Section 2, where we had only

Table 1: IGM concerns with Rows vs. Equality of Opportunity with Columns

	Circumstances	Merit/Ability
Hereditary	Σ_{-}^{1}, e.g., Race	Σ_{+}^{1}, e.g., Height
Idiosyncratic	Σ_{-}^{0}, e.g., Gender	Σ_{+}^{0}, e.g., Genius

Notes: This table presents a classification of the exogenous parameters influencing an individual's decision-making process. The examples provided within each cell are illustrative and intended for expository purposes; they should be interpreted with caution.

two exogenous parameters, $\sigma^1 = B$ and $\sigma^2 = \theta$.[12]

To talk about EOp, we also need to partition the exogenous parameters according to whether they are "morally arbitrary" for which we believe individuals should not be held accountable (circumstances) or not (abilities), $I = \Sigma_{-} \cup \Sigma_{+}$. This partition between Σ_{-} and Σ_{+} is a delicate political or philosophical matter (Roemer, 1998; Dworkin, 2002; Fleurbaey, 2008). This is in sharp contrast with the first partition between Σ^1 and Σ^0, which is statistical. We summarize these partitions of I described in the table below.

IGM of outcome implies zero correlation between outcomes of generations, that is $Y\left(x^{*}\left(\widetilde{\sigma}\right), \widetilde{\sigma}\right)$ and $Y\left(x^{*}\left(\sigma\right), \sigma\right)$. In comparison, IGM of welfare requires zero correlation between utilities of generations, that is $V\left(\widetilde{\sigma}\right)$ and $V\left(\sigma\right)$.

Now, we are prepared to describe the two principal concepts of EOp. However, prior to this, we begin with an essential disclaimer. The concept of Equality of Opportunity presents a challenge in both its definition and measurement despite its inherent appeal as a notion of leveling the playing field (Becker and Chiswick, 1966; Elford, 2023). Its definition is hard as summarized by Stanford Encyclopedia of Philosophy: "Despite its familiarity and apparent popularity, the idea of Equality of Opportunity has proved at once contested and elusive (Elford, 2023). So our goal is ... This is not opportunity in the sense of (Sen, 1999), and it is still an equality of outcome not rather than equality of choice set.

EO for outcome requires equality of ability given an effort and an ability: $\frac{\partial Y}{\partial \sigma_i} = 0$ for all $i \in \Sigma_{-}$,

EO for welfare requires that welfare would be the same for those with the same ability, $\frac{\partial V}{\partial \sigma_i} = 0$ for all $i \in \Sigma_{-}$. EO for outcome requires equal outcome for equal effort, in contrast to EO requiring equal outcome for equal cost.[13] Having established a definition for these concepts, we now turn our attention to exploring their interrelation.

8.2 Connection among Concepts

EO and IGM diverge along two dimensions.

First Dimension The difference between IGM and EO of output arises because EO holds fixed effort levels. To see this, corollary X in the Appendix, IGM of outcome is equivalent to $\frac{dY}{d\sigma_i} = 0$ for all $i \in \Sigma^1$. The idea of holding effort constant underpins the concept of EO a la Roemer and

[12]The previous framework was also more specific as it assumed linear utility and production functions, exemplified by $Y(x, \sigma) = w(\sigma)x$.

[13]To see that, write the individual problem 4 as $V\left(\sigma\right) = \max_k U\left(Y\left(K^{-1}\left(k, \sigma\right), \sigma\right), k\right)$

represents a significant hurdle in its empirical assessment, given the difficulty in observing effort.

But this discrepancy vanishes when comparing IGM and EO of welfare. To see this, first note that IGM of welfare $\frac{\partial V}{\partial \sigma_i} = 0$ for all $i \in \Sigma^1$. Then we need to appeal to the envelope theorem, implying that $\frac{\partial V}{\partial \sigma_i} = 0 \Leftrightarrow \frac{\partial Y}{\partial \sigma_i} = 0$ for any i. From a welfare perspective, the difference between IGM and EO only stems from the exogenous parameter of interest, which is the second dimension.

Second Dimension Both concepts concern about hereditary circumstance, $\Sigma^1_- = \Sigma_- \cap \Sigma^1$. In addition, EO cares about non-hereditary circumstances Σ^0_- (that is, discrimination is part of EO), whereas IGM is concerned with hereditary abilities Σ^1_+. The latter is an unintended consequence that simplifies the measurement of IGM (Intergenerational Mobility) and explains why perfect mobility is not a desired outcome.

Special case 1: Equality of opportunity=IGM of welfare when all circumstances are hereditary; no ability is hereditary, $\Sigma_- = \Sigma^1$. That is, the off-diagonal of Table 1 are empty $\Sigma^0_- = \Sigma^1_+ = \emptyset$.

Going beyond perfect equalities, IGM tends to overestimate EO since unequal opportunities can demotivate effort, leading to even larger immobility, $\Delta V \geq \Delta Y$. Mathematically, this is due to the fact that V is more convex than U,a second derivative property of an optimum value function.[14]

Special case 2: Equality of opportunity and IGM are independent concepts when all abilities are hereditary but no circumstance are hereditary, $\Sigma_- = \Sigma^0$. That is, the diagonal of Table 1 are empty $\Sigma^0_+ = \Sigma^1_- = \emptyset$.

IGM of welfare implies equality of MRS with respect to effort and hereditary features of Y and K. If we ignore the cost, $\nabla_K U = 0$, or assuming that $\nabla_\sigma K = 0$, then IGMs for welfare and outcome are equivalent. IGM of welfare $\nabla_{\sigma^1} V = 0 \Leftrightarrow \nabla_{\sigma^1} Y = 0$, using envelope. Equality of opportunity for welfare $\nabla_{\sigma_-} V = 0 \Leftrightarrow \nabla_{\sigma_-} Y = 0$, using envelope.

It is imperative to acknowledge the limitations inherent in our discussion within this subsection. Firstly, our analysis does not account for utility heterogeneity, thereby abstracting from the "expensive taste" issue. Secondly, our assumption of non-existent altruism means that the effort choices of one generation do not consider the welfare of the subsequent generation. Additionally, our discussion has omitted the consideration of "luck." We believe it is essential to clearly define the various types of luck involved to contextualize their relationship better with the elements presented in Table 1.

9 Conclusion

In 1492, Cassandra Fedele (A prominent humanist writer) wrote to Alessandra Scala (A Florentine humanist scholar): "And so, my Alessandra, you are uncertain whether to dedicate yourself to the

[14] Differentiating the Envelope condition $\frac{\partial V}{\partial \sigma_i} = \frac{\partial U}{\partial \sigma_i}$, using comparative statics: $\frac{\partial x_j}{\partial \sigma_i} = -\frac{\partial^2 U}{\partial x_j \partial \sigma_i} / \frac{\partial^2 U}{\partial x_j^2}$, gives us

$$\frac{\partial^2 V}{\partial \sigma_i^2} = \frac{\partial^2 U}{\partial \sigma_i^2} + \sum_j \frac{\partial^2 U}{\partial x_j \partial \sigma_i} \frac{\partial x_j}{\partial \sigma_i} = \frac{\partial^2 U}{\partial \sigma_i^2} + \sum_j \frac{\left(\frac{\partial^2 U}{\partial x_j \partial \sigma_i}\right)^2}{-\frac{\partial^2 U}{\partial x_j^2}} > \frac{\partial^2 U}{\partial \sigma_i^2}$$

Muses or to a Man?".[15] Both married and became successful so that they remained in recorded history. Perhaps because of their high social background.

The concept of Equality of Opportunity presents a challenge in both its definition and measurement despite its inherent appeal as a notion of leveling the playing field. Traditionally, two main approaches have emerged in economics to talk about the Equality of Opportunity. Firstly, there is the measurement of intergenerational mobility. While this approach predominantly gauges equality of outcomes, it offers the advantage of being relatively straightforward to measure. The second approach seeks to define Equality of Opportunity as achieving equal outcomes when individuals exert equal effort. This endeavor involves identifying proxies for effort. In this paper, we introduce a novel, third alternative. We propose defining Equality of Opportunity as the attainment of equal outcomes by individuals incurring equal costs. In our analysis, we quantify these costs through an examination of family investments in the available data.

[15] "Muses" refers to the Greek goddesses of inspiration in the arts, symbolizing creative or literary pursuits. See Kaborycha (2016).

References

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High wage workers and high wage firms," *Econometrica*, 67, 251–333.

ALMGREN, M., J. KRAMER, AND J. SIGURDSSON (2023): "It Runs in the Family: Occupational Choice and the Allocation of Talent," .

BEAUD, S., J. LINDGAARD, AND J. CONFAVREUX (2013): *La france invisible*, La découverte.

BECKER, G. S. AND B. R. CHISWICK (1966): "Education and the Distribution of Earnings," *The American Economic Review*, 56, 358–369.

BELL, A., R. CHETTY, X. JARAVEL, N. PETKOVA, AND J. VAN REENEN (2019): "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 134, 647–713.

BJÖRKLUND, A. AND M. JÄNTTI (1997): "Intergenerational income mobility in Sweden compared to the United States," *The American Economic Review*, 87, 1009–1018.

BJÖRKLUND, A., M. JÄNTTI, ET AL. (2009): "Intergenerational income mobility and the role of family background," in *Oxford Handbook of Economic Inequality*, Oxford University Press, vol. 491, 521.

BLACK, S. E. AND P. J. DEVEREUX (2011): "Recent developments in intergenerational mobility," *Handbook of labor economics*, 4, 1487–1541.

BLANDEN, J. (2013): "Cross-country rankings in intergenerational mobility: A comparison of approaches from economics and sociology," *Journal of Economic Surveys*, 27, 38–73.

CARD, D., J. HEINING, AND P. KLINE (2013): "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 128, 967–1015.

CHETTY, R. (2021): "Improving equality of opportunity: New insights from big data," *Contemporary Economic Policy*, 39, 7–41.

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," *The Quarterly Journal of Economics*, 129, 1553–1623.

CORAK, M. (2006): "Do poor children become poor adults? Lessons from a cross-country comparison of generational earnings mobility," *Emerald Group Publishing Limited*.

CORAK, M. AND P. PIRAINO (2011): "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 29, 37–68.

DWORKIN, R. (2002): *Sovereign virtue*, Harvard university press.

EDIN, P.-A., P. FREDRIKSSON, M. NYBOM, AND B. ÖCKERT (2022): "The rising return to noncognitive skill," *American Economic Journal: Applied Economics*, 14, 78–100.

ELFORD, G. (2023): "Equality of Opportunity," https://plato.stanford.edu/archives/fall2023/entries/equal-opportunity/.

ELHAKEEM, A., R. HARDY, D. BANN, R. CALEYACHETTY, T. D. COSCO, R. P. HAYHOE, S. G. MUTHURI, R. WILSON, AND R. COOPER (2017): "Intergenerational social mobility and leisure-time physical activity in adulthood: a systematic review," *Journal of Epidemiology & Community Health*, 71.

FLEURBAEY, M. (2008): *Fairness, responsibility, and welfare*, OUP Oxford.

GOLDIN, C. (1997): "Career and Family: College Women Look to the Past," *Gender and Family Issues in the Workplace*.

——— (2021): *Career and family: Women's century-long journey toward equity*, Princeton University Press.

HARRINGTON, M. (1997): *The other america*, Simon and Schuster.

HERMO, S., M. PÄÄLLYSAHO, D. SEIM, AND J. M. SHAPIRO (2022): "Labor market returns and the evolution of cognitive skills: Theory and evidence," *The Quarterly Journal of Economics*, 137, 2309–2361.

JÄNTTI, M., B. BRATSBERG, K. ROED, O. RAAUM, R. NAYLOR, E. OSTERBACKA, A. BJORKLUND, AND T. ERIKSSON (2006): "American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States," .

JONES, C. I. (2022): "The past and future of economic growth: A semi-endogenous perspective," *Annual Review of Economics*, 14, 125–152.

KABORYCHA, L., ed. (2016): *Corresponding Renaissance: Letters Written by Italian Women, 1375-1650*, Oxford: Oxford University Press, edited and translated by Lisa Kaborycha.

MAJOR, L. E. AND S. MACHIN (2018): *Social mobility: And its enemies*, Penguin UK.

MOGSTAD, M. AND G. TORSVIK (2023): "Family background, neighborhoods, and intergenerational mobility," *Handbook of the Economics of the Family*, 1, 327–387.

NAKAMURA, E., J. SIGURDSSON, AND J. STEINSSON (2022): "The gift of moving: Intergenerational consequences of a mobility shock," *The Review of Economic Studies*, 89, 1557–1592.

NICO, M. (2021): "Identity and change of a field: A literature analysis of the concept of social mobility," *Social Science Information*, 60, 457–478.

PUTNAM, R. D. (2016): *Our kids: The American dream in crisis*, Simon and Schuster.

ROEMER, J. E. (1998): "Equality of Opportunity," .

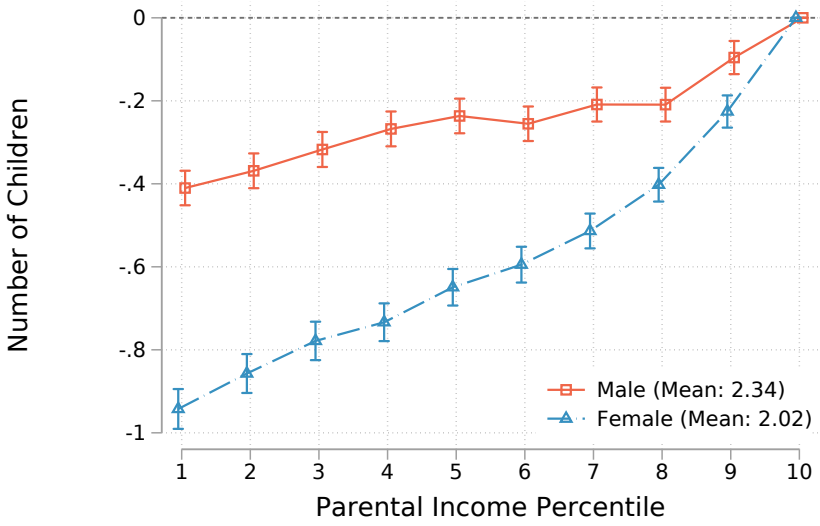
SEN, A. (1999): *Commodities and Capabilities: Amartya Sen*, Oxford University Press.

SOLON, G. (1999): "Intergenerational mobility in the labor market," in *Handbook of labor economics*, Elsevier, vol. 3, 1761–1800.

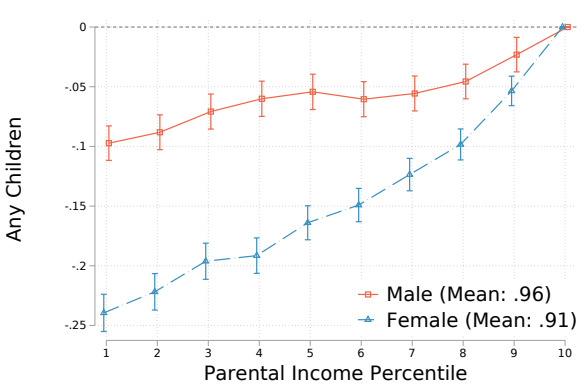
SOROKIN, P. A. (1927): *Social and Cultural Mobility*, The Free Press of Glencoe, Illinois.

VAN LEEUWEN, M. H. AND I. MAAS (2010): "Historical studies of social mobility and stratification," *Annual Review of Sociology*, 36, 429–451.

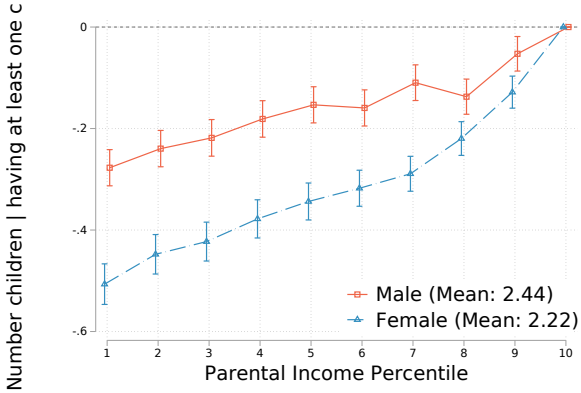
Figure 3: Family Cost of Upward Mobility



(a) Lifetime Number of Children



(b) Extensive Margin



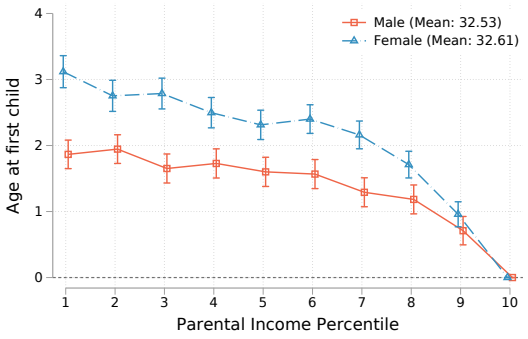
(c) Intensive margin

Note: This figure illustrates the difference in family outcomes between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. It plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of family outcome K_i of individual i on a gender-specific non-parametric functions of their own lifetime income C_i and their parental lifetime income B_i:

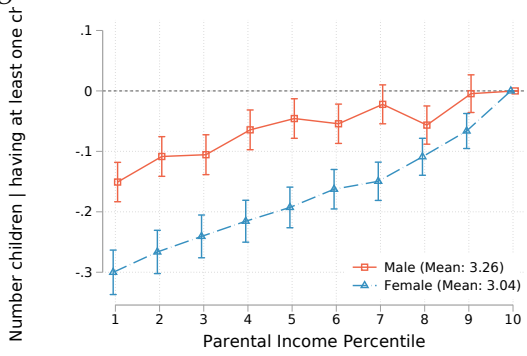
$$K_i = f^{s_i}(C_i) + g^{s_i}(B_i) + \sum_{l=1}^{9} \beta_l \underbrace{C_i}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(R_{10}(B_i) = l)}_{\text{Parental Income Deciles}} + \varepsilon_i$$

The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Own lifetime income C_i is the sum of labor incomes between ages 25 and 45. Parental lifetime income B_i is the sum of labor, business, and capital income between the individual's ages 0 and 20. Children include all biological offspring born before age 45. The functions f and g are estimated using percentile rank regression and the coefficients reported in Appendix Figure A.2. For more details, see Section 3.

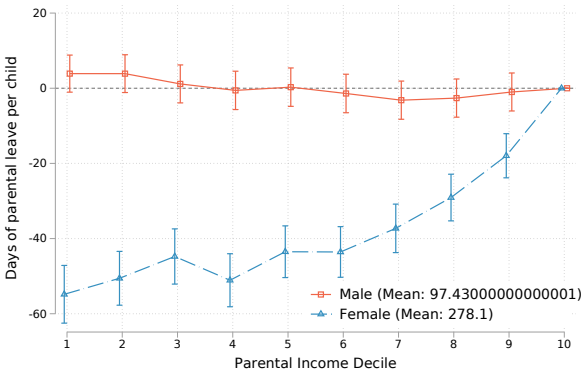
Figure 4: It's All About Time: Timing of Birth and Parental Leave



(a) Age at First Childbirth



(b) Intensive Margin conditional on age at first childbirth



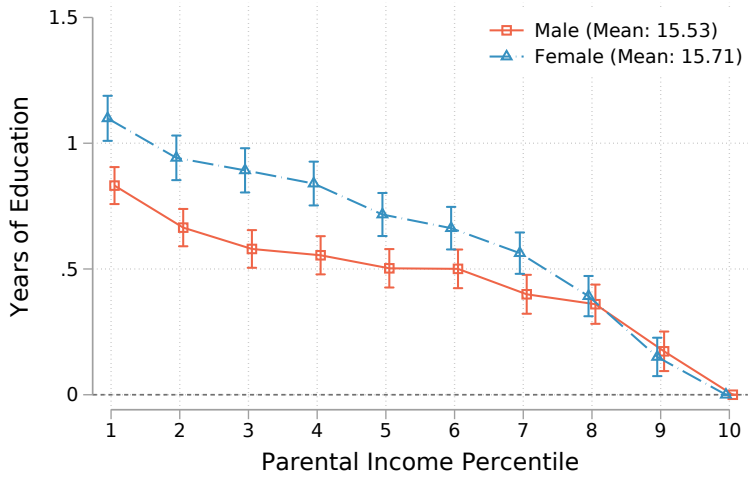
(c) Parental Leave per Child

Note: This figure illustrates the difference in family outcomes between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. It plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of family outcome K_i of individual i on a gender-specific non-parametric functions of their own lifetime income C_i and their parental lifetime income B_i:

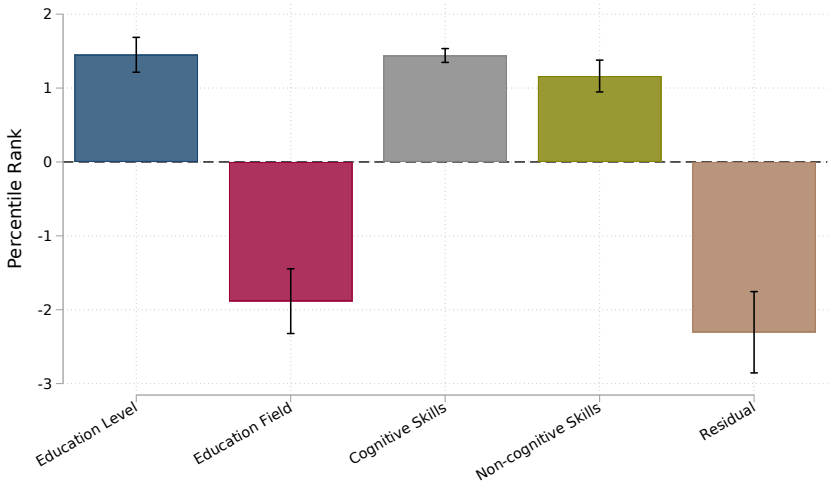
$$K_i = f^{s_i}(C_i) + g^{s_i}(B_i) + \sum_{l=1}^{9} \beta_l \underbrace{C_i}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(R_{10}(B_i) = l)}_{\text{Parental Income Deciles}} + \varepsilon_i$$

The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. All family outcomes in this figure are conditional on having at least one child. **Panel 4c** further restricts the sample to birth cohorts of 1969-1975 for whom we observe parental leave days by age 45. Own lifetime income C_i is the sum of labor incomes between ages 25 and 45. Parental lifetime income B_i is the sum of labor, business, and capital income between the individual's ages 0 and 20. Children include all biological offspring born before age 45. The functions f and g are estimated using percentile rank regression. For more details, see Section 3.

Figure 5: Mechanism: Long but Unguided Education



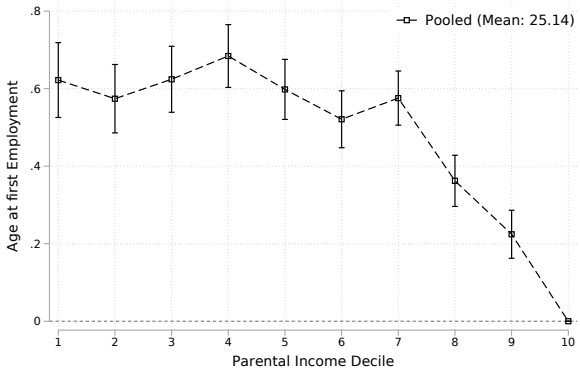
(a) Differences in Years of Education by Background Among Achievers



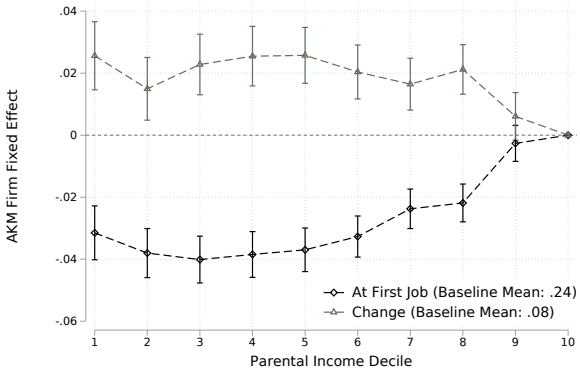
(b) Human Capital Differences by Background Among Male Achievers

Note: This figure illustrates the relative difference in human capital between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. It plots the coefficients $\{\beta_l\}$ in Regression 1 where outcomes are years of education in **Panel 5a** and four separate measures of human capital in **Panel 5b**. These four measures are harmonized using their predicted lifetime income ranks. Specifically, we use the predicted values of each measure from one regression of lifetime income rank on all human capital measures. We must restrict this exercise to men because it includes skill measures from the military draft. We report the corresponding estimates of females omitting the skill measures in Figure A.5. Cognitive skill measures are based on test scores for (i) inductive skills, (ii) verbal comprehension, (iii) spatial ability, and (iv) technical understanding. Non-cognitive skill measures stem from an interview with a certified psychologist. They include scores for (i) emotional stability, (ii) intensity, (iii) psychological energy, and (iv) social maturity. We exploit 47 distinct categories of highest completed education level-ranging from pre-school level to having a doctoral degree - and 336 categories of field of study, e.g., Teacher training, industry and crafts (Appendix Tables A.4 and A.3).

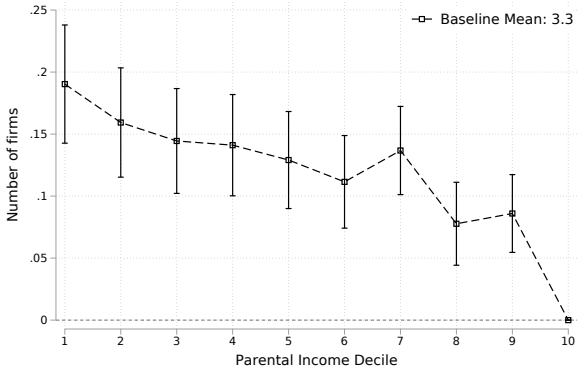
Figure 6: Achievers from Underprivilage Backgrounds are Late Bloomers



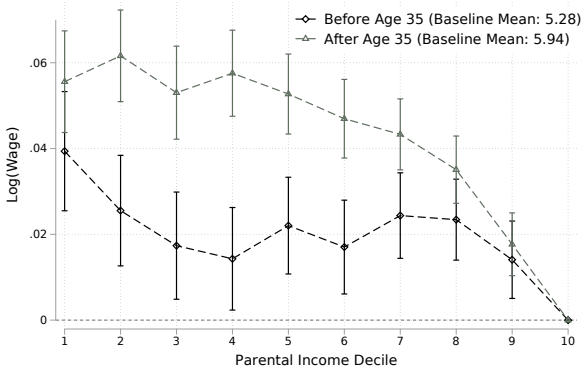
(a) Age at 1st Job



(b) AKM Firm Fixed Effect



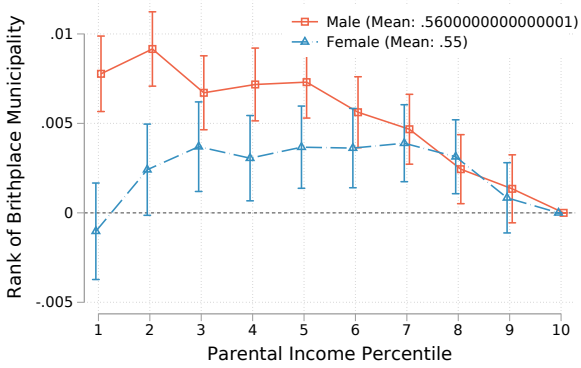
(c) Number of Jobs



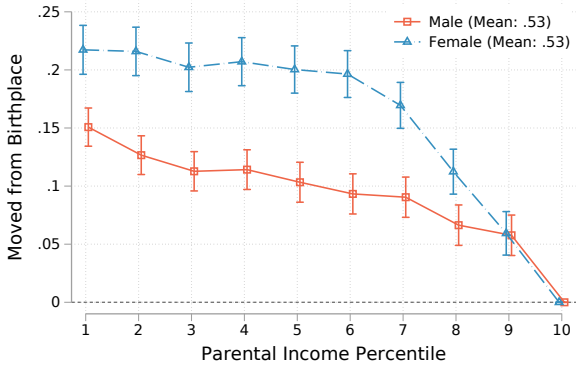
(d) Average Wage

Note: This figure illustrates the relative different careers between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group ($\{\beta_l\}$ of Equation 1). The outcome of **Panel 6b** illustrates the trajectory of the AKM firm fixed effect, including the fixed effect at the first job and the change of the fixed effect at age 45 relative to the first job. The AKM firm fixed effect captures the firm pay premium in the fashion of Abowd et al. (1999); Card et al. (2013). **Panel 6c** captures the total number of different firms between ages 30 and 45, due to the employer-employee dataset starting in 1985 when our oldest cohort was 30. **Panel 6d** reports the average wage before and after age 35. The estimation sample is based on 1,563,280 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The number is slightly below the main sample specification as we exclude individuals who have missing firm identifiers. **Panel 6d** further restricts the sample to individuals from the wage survey born between 1960 and 1975 due to the survey starting in the year 1995. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details, see Section 3.

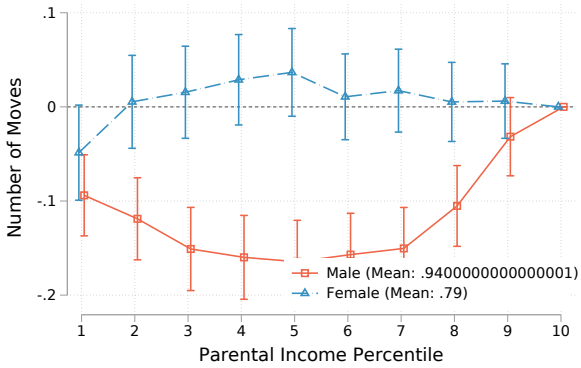
Figure 7: Geographical Location



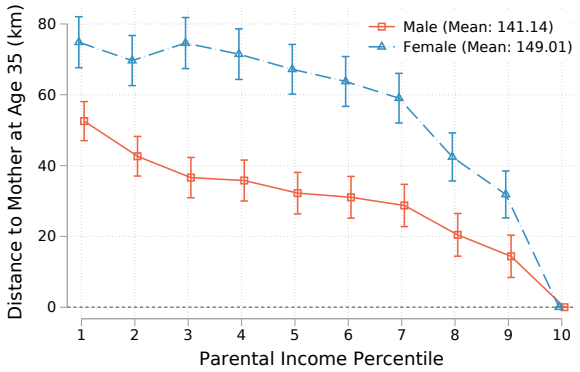
(a) Birthplace Income Rank



(b) Moved From Birthplace



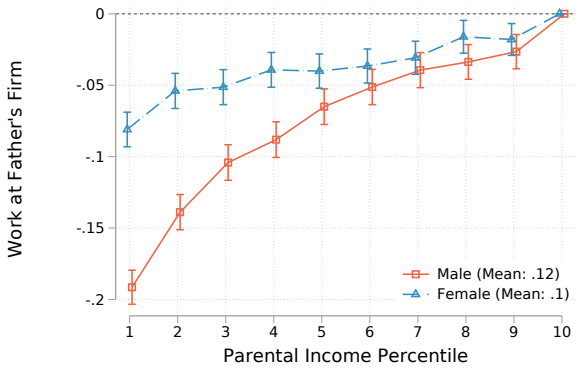
(c) Number of Moves



(d) Distance to Mother

Note: This figure illustrates the relative different careers between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group ($\{\beta_l\}$ of Equation 1). The outcome of **Panel 7a** measures the average population income rank of working-age individuals living in the municipality at the year of birth. **Panel 7b** describes an indicator that takes the value one if the individual's county of residence at age 45 differs from the county of birth. In **Panel 7c**, we measure the dependent variable by counting the number of residence changes between the ages of 35 to 45. Finally, **Panel 7d** reports the straight-line distance in km between the individual's residence address and the mother's residence address. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details, see Section 3.

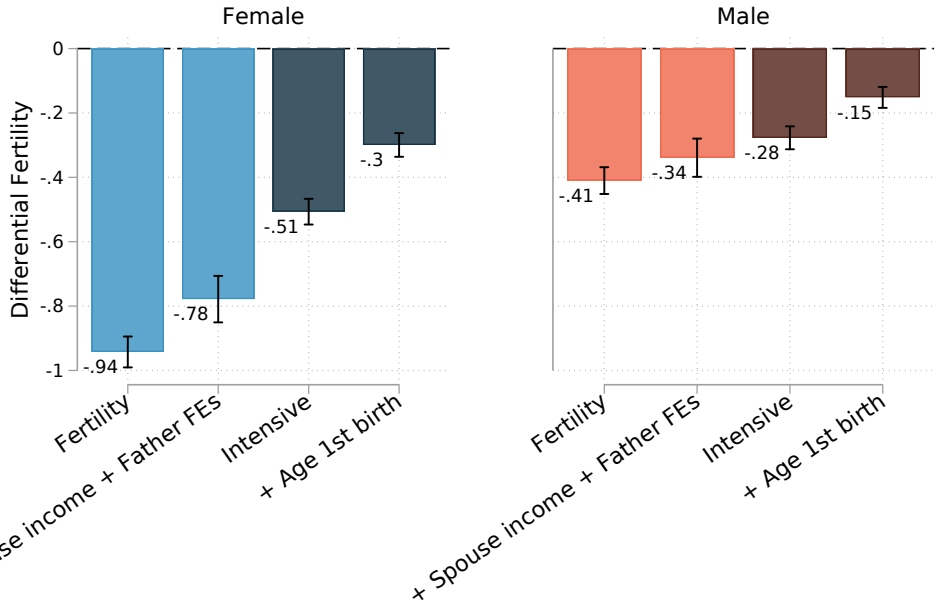
Figure 8: Networks



(a) Same Firm as Father

Note: This figure illustrates the relative difference careers between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group (β_l of Equation 1). In **Panel 8a**, we measure the outcome as working in a firm between ages 30 and 45 that the father has ever worked for. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details, see Section 3.

Figure 9: Adding controls



Note: This figure shows the relative difference in the number of children between individuals at the top vs. the bottom of lifetime income for individuals with bottom vs. top parental income, where top decile parents are the baseline group. It plots the coefficients $\{\beta\}$ in Regression 2 where we additionally add control variables to the main regression. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden.

Online Appendix of:

Invisible Cost of Upward Mobility

Patrizia Massner & Arash Nekoei

March 20, 2025

A Theory

In this appendix we provide the proves of all claims in the main text, first for the case where γ is completely hereditary, $\gamma = \widetilde{\gamma}$, and then for the general case. The CDF of random variable X is denoted by $F_X$and $\mathbb{1}$ denotes the indicator function. To fix ideas, we can sue this notation and write: $F_X(x) = \mathbb{P}(X < x) = \mathbb{E}(\mathbb{1}(X < x))$.

Mobility Definition We measure mobility as negative of Kendall's rank correlation of parent and child outcome. Kendall's rank correlation is a measure of dependence between two random variable. It is defined as the probability of concordances minus discordances.. \cite{schweizer1981}

$$\tau\left(Z, \widetilde{Z}\right) = \underbrace{\mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) > 0\right)}_{\text{Prob of concordances}} - \underbrace{\mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) < 0\right)}_{\text{Prob of discordances}}$$

So $\tau\left(Z, \widetilde{Z}\right) \in [-1, 1]$. We can also write the coefficient using the expected value of the copula

$$\tau\left(Z, \widetilde{Z}\right) = 4\mathbb{E}\left(C_{Z,\widetilde{Z}}\left(Z, \widetilde{Z}\right)\right) - 1 \tag{5}$$

To see this, we can rewrite the Kendall's rank correlation as a function of only the probability of concordances

$$\begin{aligned}\tau\left(Z, \widetilde{Z}\right) &= \mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) > 0\right) - P\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) < 0\right) \\ &= 2\mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) > 0\right) - 1\end{aligned} \tag{6}$$

We then use the symmetry between $\left(Z_1, \widetilde{Z}_1\right)$ and $\left(Z_2, \widetilde{Z}_2\right)$ to write the Kendall's rank correlation as

$$\tau\left(Z, \widetilde{Z}\right) = 4\mathbb{P}\left(Z_1 < Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right) - 1 \tag{7}$$

Now we can connect the right hand side to the joint CDF and copula using the law of iterated

expectations:

$$\begin{aligned}\mathbb{P}\left(Z_1 < Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right) &= \mathbb{E}\left(\mathbb{1}\left(Z_1 < Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right)\right) \\ &= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}\left(Z_1 < Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right)|Z_2, \widetilde{Z}_2\right)\right) \\ &= \mathbb{E}\left(F_{Z,\widetilde{Z}}\left(Z, \widetilde{Z}\right)\right) \\ &= \mathbb{E}\left(C_{Z,\widetilde{Z}}\left(Z, \widetilde{Z}\right)\right)\end{aligned}$$

Expressions 7 and 5 are useful in practice. In particular, using the former, we can write the Kendall's rank correlation as a decreasing function of the half-probability of discordances, that is

$$\tau\left(Z, \widetilde{Z}\right) = 1 - 4\chi\left(Z, \widetilde{Z}\right) \tag{8}$$

where $\chi\left(Z, \widetilde{Z}\right) = \mathbb{P}\left(Z_1 > Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right)$ is the half probability of discordances. This implies that our mobility definition is an increasing function of the half probability of discordances. In this sense, mobility is the likelihood of two families interchange their relative position across generations.

Hereditary Process In the main text, we assumed that γ is completely hereditary, $\gamma = \widetilde{\gamma}$. In general case includes the case where the hereditary factor is the prior productivity itself, $\gamma_i = \widetilde{w}_i$, or equivalently productivities are AR1 case of

$$\ln w_i = \alpha_i \mu_i + \alpha_i \ln \theta + (1 - \alpha_i) \ln \widetilde{w}_i$$

For comparison, the data generating process in the main text is

$$\ln w_i = \ln A_i + (1 - \alpha_i) \ln \gamma + \alpha_i \ln \theta$$

Proof. of Proposition **??**

Individuals allocate their time to balance the personal and professional life, $t_c = \left[\frac{A_c}{A_k}\left(\frac{\theta}{\gamma}\right)^{\alpha_c - \alpha_k}\right]^{\varepsilon}$. This leads to the optimum career of $C = \frac{A_c^{\varepsilon} A_k^{1-\varepsilon}}{1-\frac{1}{\varepsilon}} \theta^m \gamma^{1-m}$, where $m = \varepsilon\alpha_c + (1-\varepsilon)\alpha_k$ is the weighted average of meritocracy level in career and family. The optimal family outcome can be written as

$$\begin{aligned}K &= A_k \theta^{\alpha_k} \gamma^{1-\alpha_k} - A_c A_k^{1-\varepsilon} \theta^m \gamma^{1-m} \\ &= \left(\frac{A_k^{\alpha_c}}{A_c^{\alpha_k}}\right)^{\frac{\varepsilon}{m}} \gamma^{1-\frac{\alpha_k}{m}} \left[\left(1 - \frac{1}{\varepsilon}\right) C\right]^{\alpha_k/m} - \left(1 - \frac{1}{\varepsilon}\right) C\end{aligned}$$

This implies that $\frac{d}{d\gamma}\frac{dK}{dC}$ has the same sign than $\left(1 - \frac{\alpha_k}{m}\right)\frac{\alpha_k}{m}$. So $\frac{d}{d\gamma}\frac{dK}{dC} < 0$ iff $\frac{\alpha_k}{m} > 1$. $Q.E.D.$

Extension with AR1, $\frac{d}{d\gamma_{-1}}\mathbb{E}_{\Theta}\left(\frac{dK}{dC}\right) < 0$ which means conditioning on both parental productiv-

ities, $\widetilde{w}_c$ and $\widetilde{w}_k$.

Ranking of Outcomes by their Mobility The probability of having higher outcome as a function of hereditary features of individuals is denoted by

$$\phi_g(\gamma_1, \gamma_2) = \mathbb{P}(g(\gamma_2, \Theta_2) > g(\gamma_1, \Theta_1))$$

where Θ_1 and Θ_2 are IID with distribution Θ. Similarly, the probability of having higher outcome as a function of hereditary potential of individuals is denoted by

$$\widetilde{\phi}_g(\widetilde{\gamma}_1, \widetilde{\gamma}_2) = \mathbb{P}(\phi_g(\Gamma_1, \Gamma_2) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2)$$

where Θ_1 and Θ_2 are IID with distribution Θ and $\Gamma_i|\widetilde{\gamma}_i$ determined by our assumption of data generating process of hereditary factor. Note that the former probability does not depend on the hereditary process. The two probabilities are the same under the fully hereditary assumption in the main text.

For the outcome function g, the mobility is measured as the negative of the rank correlation between outcomes of two generations, $-\tau_g$. The rank correlation τ_g can be written as the function of the half probability of discordances by χ_g, using expression 8, that is

$$\tau_g = \tau\left(g(\Gamma, \Theta), g\left(\widetilde{\Gamma}, \widetilde{\Theta}\right)\right) = 1 - 4\,\chi_g \tag{9}$$

where Θ and $\widetilde{\Theta}$ are IID with distribution Θ and the joint distribution of Γ and $\widetilde{\Gamma}$ is determined by our assumption of data generating process of hereditary factor.

Definition 1. *The output function g is more mobile than outcome h when Kendall's τ is lower for the former, that is, $\tau_h > \tau_g$, or equivalently, $\chi_h > \chi_g$.*

A stronger alternative but a non-complete ranking of outcomes by their mobility is that any agent with a lower family background has a always higher chance of taking over her privileged peer. However, this does not mean that A's parent *outcome* has been lower than B's parent, since A's parent might have been much more talented than B's parent so that they undone their under-privilege.

Definition 2. *The output g is strictly more mobile than outcome h when for all $(\widetilde{\gamma}_1, \widetilde{\gamma}_2) \in \mathbb{R}^2_+$,*

$$\widetilde{\gamma}_1 > \widetilde{\gamma}_2 \;\Rightarrow\; \widetilde{\phi}_g(\widetilde{\gamma}_1, \widetilde{\gamma}_2) > \widetilde{\phi}_h(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$$

This definition is suited for the model with fully hereditary feature We also show that strict mobility ranking is not a complete ranking and it implies non-strict mobility ranking, which is complete. The difference in completeness is due to the fact that strict mobility requires higher rank breaking in all cases, whereas the new requirement is in expectation.

A stronger version of the strict mobility above is when for all $(\widetilde{\gamma}_1, \widetilde{\gamma}_2) \in \mathbb{R}^2_+$,

$$\widetilde{\gamma}_1 > \widetilde{\gamma}_2 \;\; \Rightarrow \;\; \phi_g\left(\widetilde{\gamma}_1, \widetilde{\gamma}_2\right) \; > \; \phi_h\left(\widetilde{\gamma}_1, \widetilde{\gamma}_2\right)$$

The latter implies the former and they are equivalent in the case of fully hereditary.

One additional definition is useful before delving into our claims and propositions.

Claim 1. *The probability of discordances for outcome g can be written as* $\mathbb{P}\left(\frac{1}{2}\left(\Phi_g + \widetilde{\Phi}_g\right) - \Phi_g\widetilde{\Phi}_g\right)$ *,where* $\widetilde{\Phi}_g = \mathbb{P}\left(\widetilde{\phi}_g\left(\Gamma_1, \Gamma_2\right) | \widetilde{\Gamma}_1 > \widetilde{\Gamma}_2\right)$ *and* $\Phi_g = \mathbb{P}\left(\phi_g\left(\widetilde{\Gamma}_1, \widetilde{\Gamma}_2\right) | \widetilde{\Gamma}_1 > \widetilde{\Gamma}_2\right)$ *are the measures of inter- and intra-generational mobility, respectively.*

Proof. The half probability of discordances for outcome g is defined as

$$\chi_g = \mathbb{P}\left(g\left(\Gamma_2, \Theta_2\right) < g\left(\Gamma_1, \Theta_1\right) \, , \, g\left(\widetilde{\Gamma}_2, \widetilde{\Theta}_2\right) > g\left(\widetilde{\Gamma}_1, \widetilde{\Theta}_1\right)\right)$$

We can rewrite the right hand side, first as

$$\begin{aligned}\chi_g = \mathbb{P}\Bigg(&\frac{1}{2}\mathbb{P}\left(g\left(\Gamma_2, \Theta_2\right) < g\left(\Gamma_1, \Theta_1\right) \, , \, g\left(\widetilde{\Gamma}_2, \widetilde{\Theta}_2\right) > g\left(\widetilde{\Gamma}_1, \widetilde{\Theta}_1\right) | \widetilde{\Gamma}_1 > \widetilde{\Gamma}_2\right)\Bigg) \\ + \;\; &\left(\frac{1}{2}\mathbb{P}\left(g\left(\Gamma_2, \Theta_2\right) < g\left(\Gamma_1, \Theta_1\right) \, , g\left(\widetilde{\Gamma}_2, \widetilde{\Theta}_2\right) > g\left(\widetilde{\Gamma}_1, \widetilde{\Theta}_1\right) | \widetilde{\Gamma}_1 < \widetilde{\Gamma}_2\right)\right)\end{aligned}$$

and then, given the symmetry and independence between Γ_1 and Γ_2, as

$$\begin{aligned}\chi_g = \mathbb{P}\Bigg(&\frac{1}{2}\mathbb{P}\left(g\left(\Gamma_2, \Theta_2\right) < g\left(\Gamma_1, \Theta_1\right) \, , \, g\left(\widetilde{\Gamma}_2, \widetilde{\Theta}_2\right) > g\left(\widetilde{\Gamma}_1, \widetilde{\Theta}_1\right) | \widetilde{\Gamma}_1 > \widetilde{\Gamma}_2\right)\Bigg) \\ + \;\; &\left(\frac{1}{2}\mathbb{P}\left(g\left(\Gamma_1, \Theta_1\right) < g\left(\Gamma_2, \Theta_2\right) \, , \, g\left(\widetilde{\Gamma}_1, \widetilde{\Theta}_1\right) > g\left(\widetilde{\Gamma}_2, \widetilde{\Theta}_2\right) | \widetilde{\Gamma}_1 > \widetilde{\Gamma}_2\right)\right)\end{aligned}$$

and finally, given the symmetry and independence among Θ_i and $\widetilde{\Theta}_i$, as

$$\chi_g = \mathbb{P}\left(\frac{1}{2}\left(1 - \widetilde{\Phi}_g\right)\Phi_g + \frac{1}{2}\widetilde{\Phi}_g\left(1 - \Phi_g\right)\right)$$

$Q.E.D.$

Definition 3. *For any outcome function g,we define the compensating function* $\sigma_g(.)$ *at* (a, b) *as the amount of talent* θ *necessary to compensate for lower family background at that point, that is, for any* $\Delta \in [0, a]$*,* $\sigma_g(\Delta)$ *is implicitly defined as*

$$g\left(a - \Delta, b + \sigma_g\left(\Delta\right)\right) = g\left(a, b\right)$$

Two important notes about compensating function. First, Quasi-concavity of outcome function implies diminishing MRS, \citep{arrow1961}, implying that compensating function is convex. The

second note is that the compensating function is specific to the point (a, b). This add one layer of complication to the usage of the compensating function.

Claim 2. *Hereditary feature is beneficial in expectation, that is,* $\phi_g(\gamma_1, \gamma_2) > .5 \Leftrightarrow \gamma_2 > \gamma_1$.

Proof. Using the definition of compensating function (Definition 3), we have $g(\gamma_1, \theta_1) = g(\gamma_2, \theta_1 + \sigma_g(\Delta\gamma))$. Rewrite the probability of having higher outcome as

$$\begin{aligned}\mathbb{P}(g(\gamma_2, \Theta_2) > g(\gamma_1, \Theta_1)) &= \mathbb{P}(\mathbb{P}(g(\gamma_2, \Theta_2) > g(\gamma_1, \Theta_1) \mid \Theta_1)) \\ &= \mathbb{P}(\mathbb{P}(\sigma_g(\gamma_2 - \gamma_1) > \Delta\Theta \mid \Theta_1)) > \mathbb{P}(\mathbb{P}(0 > \Delta\Theta \mid \Theta_1)) \\ &= \mathbb{P}(0 > \Delta\Theta) = .5\end{aligned}$$

where $\Delta\Theta = \Theta_1 - \Theta_2$ using the fact that Θ_i are IID.

$$\begin{aligned}\widetilde{\phi}_g(\widetilde{\gamma}_1, \widetilde{\gamma}_2) &= \mathbb{P}(\phi_g(\Gamma_1, \Gamma_2) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2) \\ &= \mathbb{P}(\mathbb{P}(\sigma_g(\Delta\Gamma) > \Delta\Theta) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2) \\ &= \mathbb{P}(\mathbb{P}(\sigma_g(\Delta\Gamma) > \Delta\Theta) + \mathbb{P}(\sigma_g(\Delta\Gamma) > \Delta\Theta) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2, \Gamma_1 > \Gamma_2)\end{aligned}$$

where $\Delta\Theta = \Theta_1 - \Theta_2$ and $\Delta\Gamma = \Gamma_2 - \Gamma_1$. Convexity of compensating function implies that $\sigma_g(x) + \sigma_g(-x) > 0$ $Q.E.D.$

Claim 3. *The strict higher mobility implies higher mobility.*

Proof. Claim 2 implies that$\Phi_g < .5$ and $\widetilde{\Phi}_g < .5$ [not true].We then use claims 1, and Strict higher mobility of g means that Φ_g is larger than Φ_h. $Q.E.D.$

Proof. of Lemma **??**

I. Proof for the case of full hereditary, $\gamma = \widetilde{\gamma}$.

We first use the definition of compensating function (Definition 3) and show that $\frac{\partial \sigma_g}{\partial \gamma} = \frac{\partial g}{\partial \gamma} / \frac{\partial g}{\partial \theta}$ and $\sigma_g(0) = 0$.

The higher MRS of g thus implies that $\sigma_g(\gamma) > \sigma_h(\gamma)$ when $\gamma > 0$ and $\sigma_g(\gamma) < \sigma_h(\gamma)$ when $\gamma < 0$. Now again using the definition of compensating function, we have that

$$g(\gamma_1, \Theta_1) = g(\gamma_2, \Theta_1 + \sigma_g(\Delta\gamma)) < g(\gamma_2, \Theta_2)$$

where $\Delta\gamma = \gamma_1 - \gamma_2 > 0$. Given the outcome function g is increasing, we have that

$$g(\gamma_2, \Theta_2) > g(\gamma_1, \Theta_1) \Leftrightarrow \sigma_g(\Delta\gamma) < \Delta\Theta$$

where $\Delta\Theta = \Theta_2 - \Theta_1$. Applying this equivalence to both outcome functions, we get

$$\phi_g(\gamma_1, \gamma_2) = P(\sigma_g(\Delta\gamma) < \Delta\Theta)$$

and

$$\phi_h(\gamma_1, \gamma_2) = P(\sigma_h(\Delta\gamma) < \Delta\Theta)$$

Together, they imply that

$$\phi_g(\gamma_1, \gamma_2) < \phi_h(\gamma_1, \gamma_2) \Leftrightarrow \sigma_h(\Delta\gamma) < \sigma_g(\Delta\gamma)$$

II. To extend the proof to the case where γ has the Markov property, we use the law of iterated expectations,

$$\begin{aligned}
\widetilde{\phi}_g(\widetilde{\gamma}_1, \widetilde{\gamma}_2) &= \mathbb{P}(g(\Gamma_2, \Theta_2|\widetilde{\gamma}_2) > g(\Gamma_1, \Theta_1|\widetilde{\gamma}_1)) \\
&= \mathbb{E}(\mathbb{1}(g(\Gamma_2, \Theta_2) > g(\Gamma_1, \Theta_1)) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}(g(\Gamma_2, \Theta_2) > g(\Gamma_1, \Theta_1)) | \Gamma_1, \Gamma_2) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}(\sigma_g(\Delta\gamma) < \Theta_2 - \Theta_1) | \Gamma_1, \Gamma_2) \mid \widetilde{\gamma}_1, \widetilde{\gamma}_2)
\end{aligned}$$

$Q.E.D.$

Proof. of Proposition **??**

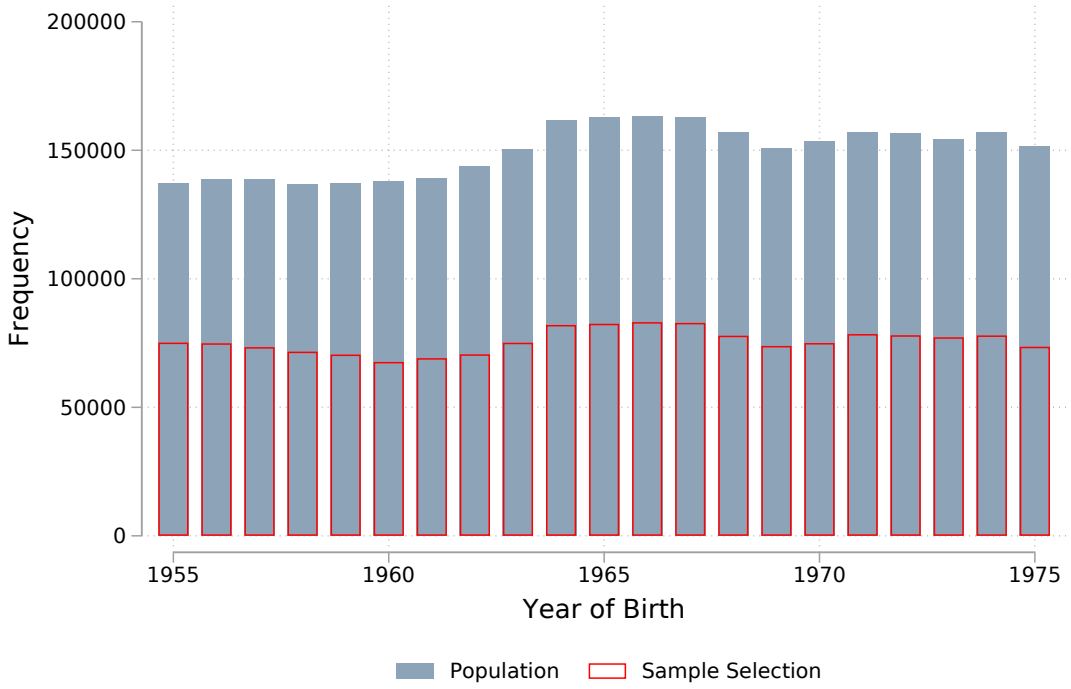
Proposition **??** implies that $\frac{d}{d\gamma}\frac{dK/d\theta}{dC/d\theta} < 0 \Leftrightarrow \alpha_k > \alpha_c$. We show below $\alpha_k > \alpha_c$ is equivalent to higher MRS for income than welfare. To show that $\alpha_k > \alpha_c$ is equivalent to higher MRS for income than welfare, start with the optimal welfare is $U = A_k\theta^{\alpha_k}\gamma^{1-\alpha_k} + \frac{A_c A_k^{1-\varepsilon}}{\varepsilon-1}\theta^m\gamma^{1-m}$ for individual (γ, θ), where $m = \varepsilon\alpha_c + (1-\varepsilon)\alpha_k$ is again the weighted average of meritocracy level in career and family. . So $\frac{dC}{d\gamma}/\frac{dC}{d\theta} > \frac{dU}{d\gamma}/\frac{dU}{d\theta}$ can be simplified as

$$\frac{1-m}{\gamma}/\frac{m}{\theta} > \frac{(1-\alpha_k)K + (1-m)C}{\gamma}/\frac{\alpha_k K + mC}{\theta}$$

which is equivalent to $\frac{1-m}{m} > \frac{1-\alpha_k}{\alpha_k}$ and $\alpha_k > \alpha_c$. To complete the proof, we use lemma **??** that showed that higher MRS implies strong strict mobility order. $Q.E.D.$

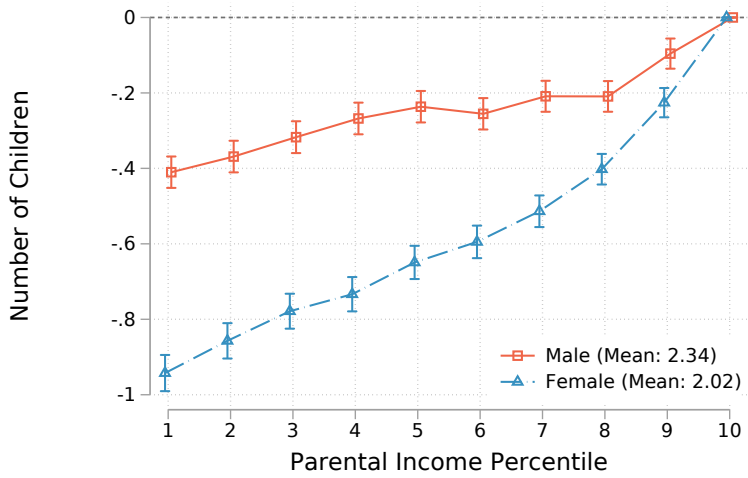
B Appendix Figures

Figure A.1: Number of Individuals in the Main Sample

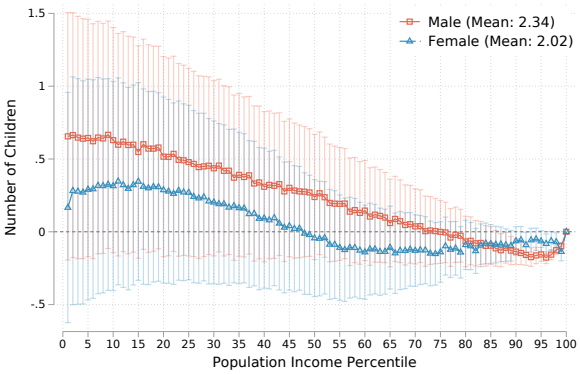


Note: This figures displays the number of individuals in our main sample selection against the population of Sweden born in the specific birth cohort. Our main sample consists of is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden.

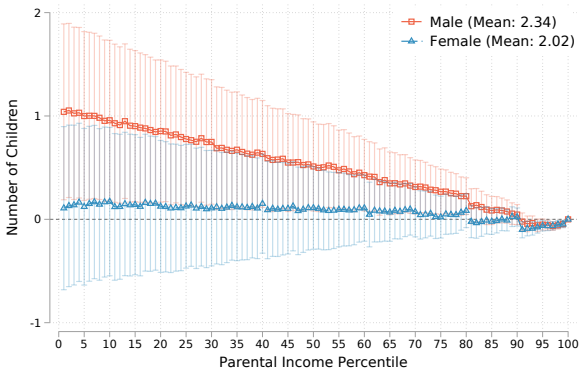
Figure A.2: Cost of Mobility measured as the Number of Children



(a) Cost of Mobility: Interaction, β_l



(b) Own Income Pct Fixed Effects, δ_l



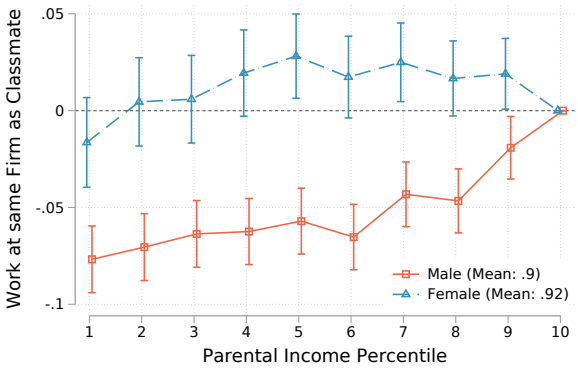
(c) Parental Income Pct Fixed Effects, γ_l

Note: **Panel A.2a** illustrates the relative difference in family outcomes between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. It plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of family outcome K_i of individual i on a gender-specific non-parametric functions of their own lifetime income C_i and their parental lifetime income B_i:

$$K_i = f^{s_i}(C_i) + g^{s_i}(B_i) + \sum_{l=1}^{9} \beta_l \underbrace{C_i}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(R_{10}(B_i) = l)}_{\text{Parental Income Deciles}} + \varepsilon_i$$

The outcome variable K_i captures the number of children before the age of 45. **Panel A.2b** and **Panel A.2c** display the corresponding estimates of own lifetime income and parental lifetime income, respectively. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Own lifetime income C_i is the sum of labor incomes between ages 25 and 45. Parental lifetime income B_i is the sum of labor, business, and capital income between the individual's ages 0 and 20. Children include all biological offspring born before age 45. For more details see Section 3.

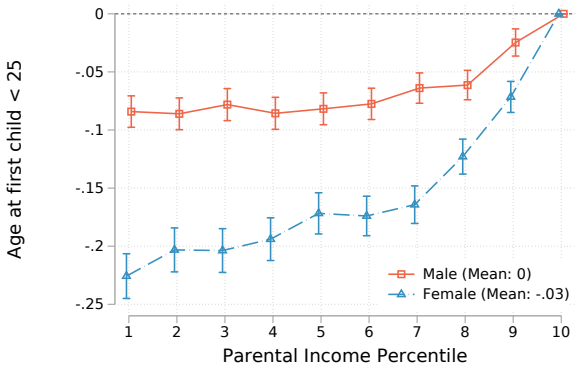
Figure A.3: Classmate Networks



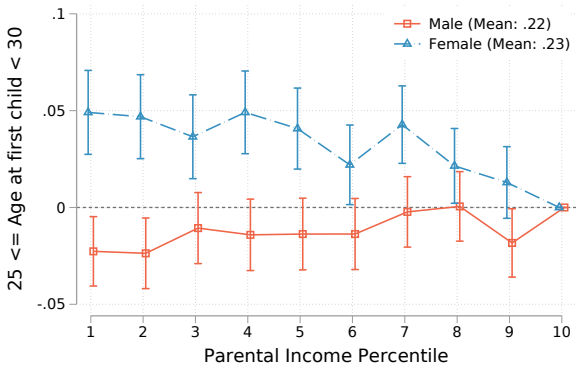
(a) Same Firm as Classmate Network

Note: This figure illustrates the relative difference in family outcomes between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group (β_l of Equation 1). In **Panel A.3a**, we measure the outcome as having worked in the same firm between ages 30 and 45 as any firm their university peers have ever worked for. We define university peers as the group of individuals that graduate in the same year, the same university major, and the same municipality. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details, see Section 3.

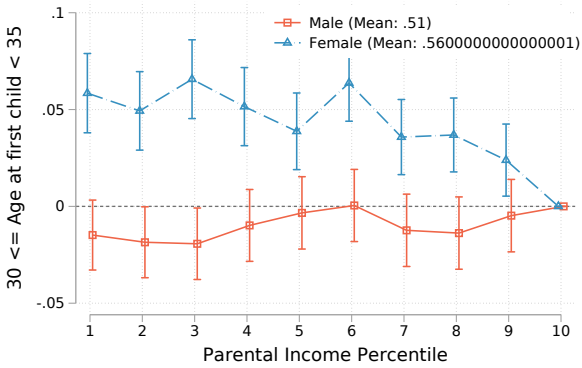
Figure A.4: Age at First Birth, Non-parametric Analysis



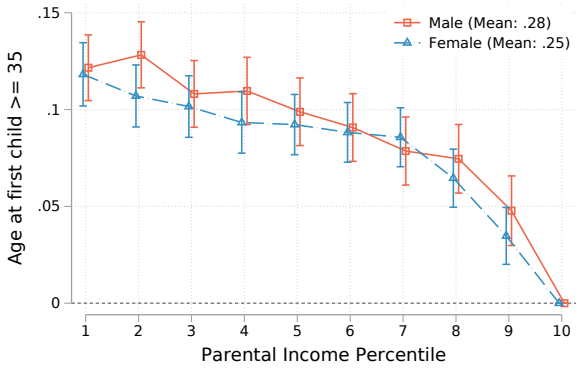
(a) Age at First Birth below 25



(b) Age at First Birth between 25 and 30



(c) Age at First Birth between 30 and 35



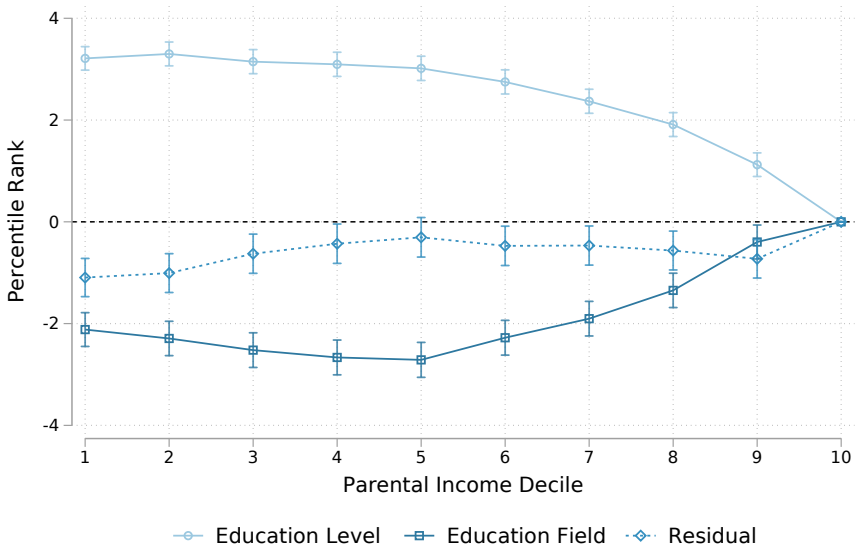
(d) Age at First Birth above 35

Note: This figure illustrates the relative difference in family outcomes between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. It plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of family outcome K_i of individual i on a gender-specific non-parametric functions of their own lifetime income C_i and their parental lifetime income B_i:

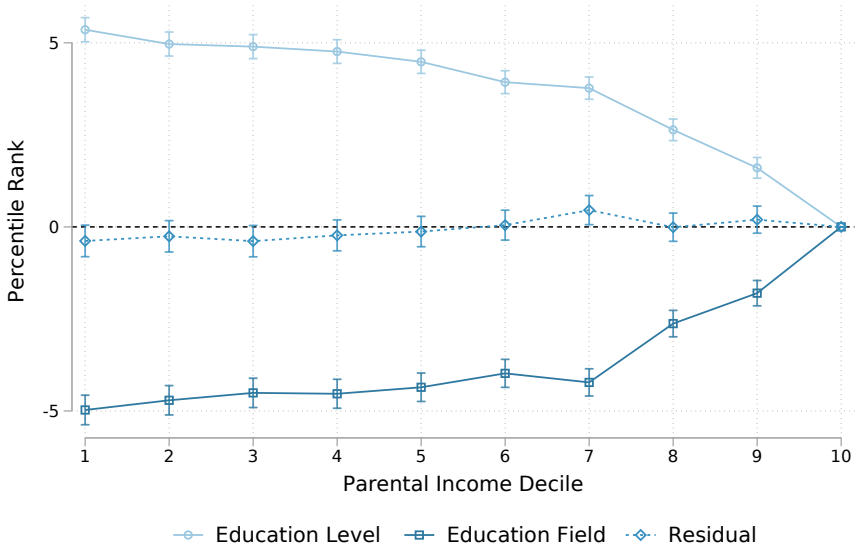
$$K_i = f^{s_i}(C_i) + g^{s_i}(B_i) + \sum_{l=1}^{9} \beta_l \underbrace{C_i}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(R_{10}(B_i) = l)}_{\text{Parental Income Deciles}} + \varepsilon_i$$

The outcome variables K_i capture the individual's age group at first birth before the age of 45 conditional on having at least one child during that time. For example, **Panel A.4a** represents an indicator that takes the value of one of the age at first birth is below the age of 25. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is 1,349,653 individuals with at least one child born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details see Section 3.

Figure A.5: Diverse Human Capital Paths to the Top by Background



(a) Male



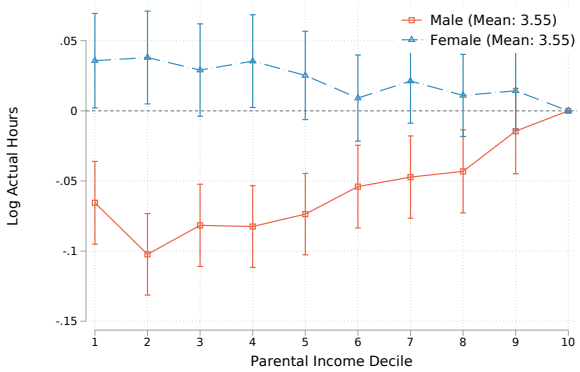
(b) Female

Note: This figure illustrates the relative difference in human capital between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group, $\{\beta_l\}$ in Regression 1. We use two separate measures of human capital: education level and education field (see Table A.3 and A.4). They are harmonized using their respective predicted lifetime income ranks. For this, we first run the following non-parametric regression of lifetime income rank on all human capital measures:

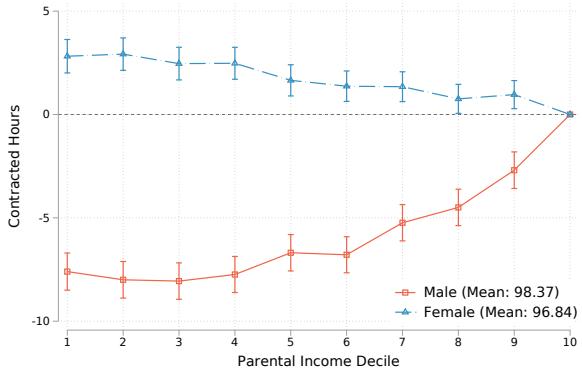
$$\underbrace{R(C_i)}_{\text{Own LT Income Rank}} = \underbrace{\xi_k}_{\text{Edu Level FE}} + \underbrace{\zeta_j}_{\text{Edu Field FE}} + \epsilon_i$$

where we run the regression separately by gender g. The R^2 of the regression is 0.214 (male) and 0.247 (female). We construct the outcome variables as the predicted values of the respective human capital measure and use them in the main specification of Equation 1. **Panel A.5a** and **Panel A.5b** report the respective interaction terms $\{\beta_l\}$. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. For more details see Section 3.

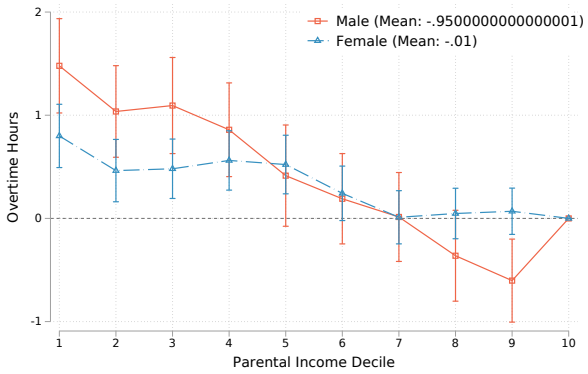
Figure A.6: Hours Worked



(a) Actual Hours Worked



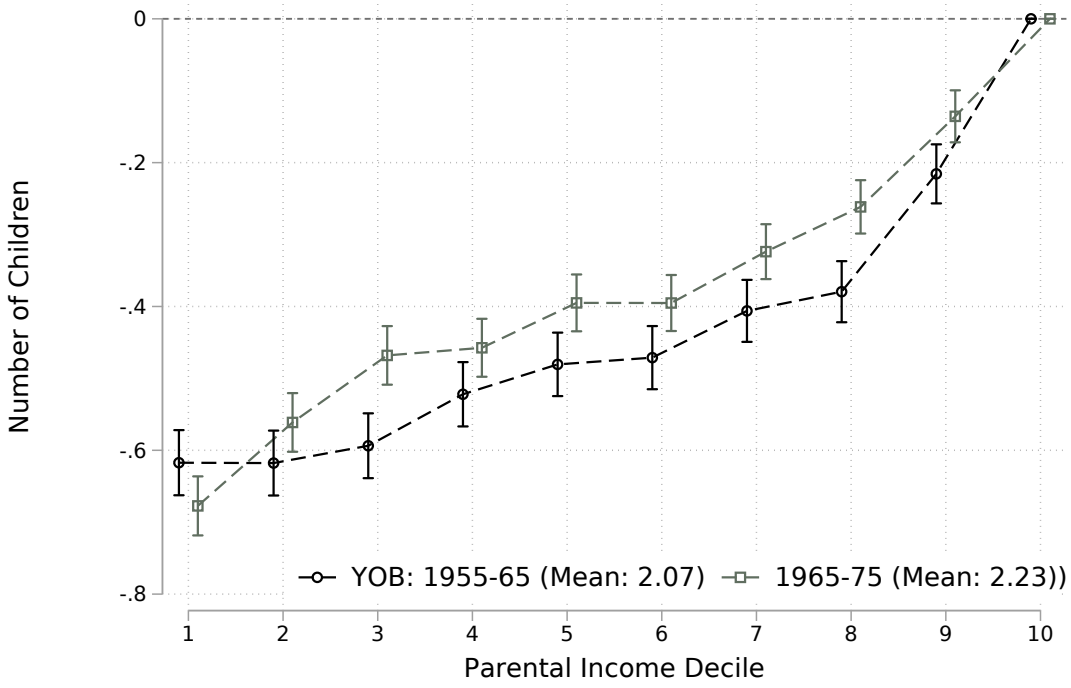
(b) Contracted Hours



(c) Overtime Hours

Note: This figure illustrates the relative different hours worked between individuals at the top vs. the bottom of lifetime income by parental income deciles, using top-decile parents as the reference group. ($\{\beta_l\}$ of Equation 1). The outcome of **Panel A.6a** measures the log of monthly actual hours worked. **Panel A.6b** describes contracted hours worked measured as the share of full-time work. **Panel A.6b** describes the monthly hours worked in overtime. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. We only consider individuals covered by the Wage Structure Survey for this exercise. The survey covers a representative sample of the Swedish population in the labor force where monthly hours worked are reported by the employer as of September of each year. The baseline mean reports the average of dependent variables among the top-earning individuals with top-earning parents. For more details, see Section 3.

Figure A.7: Evolution of Cost of Mobility over Time

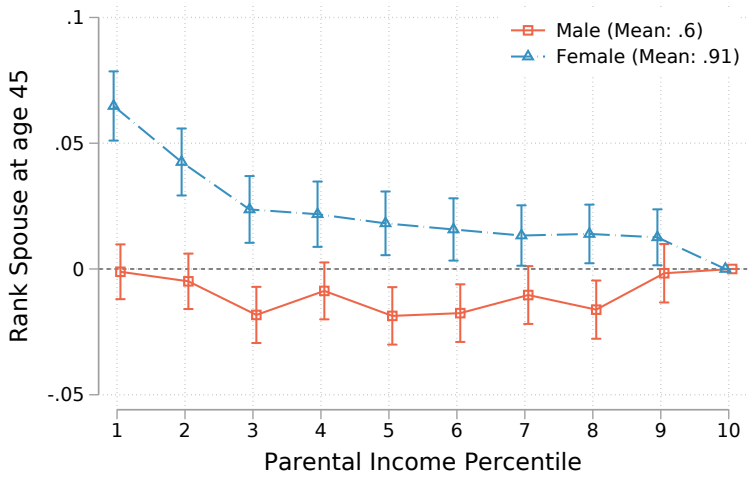


Note: This figure shows the coefficient estimates of running the following regression separately by birth cohort groups:

$$K_i = f^{s_i}(C_i) + g^{s_i}(B_i) + \sum_{l=1}^{9} \beta_l \underbrace{C_i}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(R_{10}(B_i) = l)}_{\text{Parental Income Deciles}} + \varepsilon_i$$

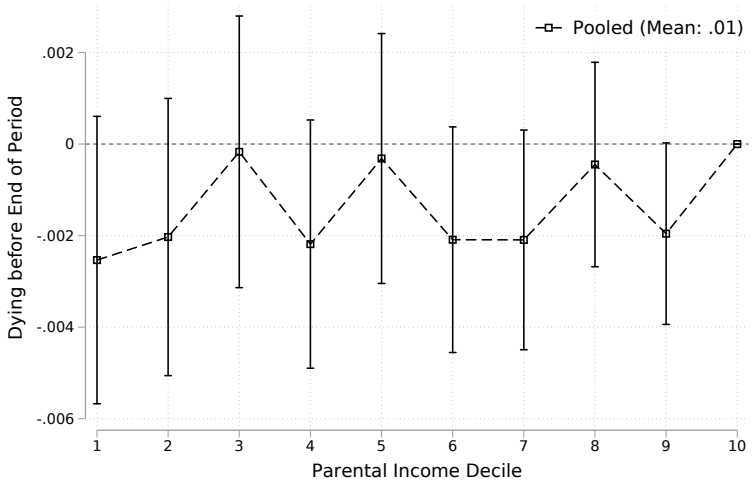
we run the regression pooled by gender but interact control variables by gender. We categorize individuals into two groups based on their year of birth (YOB): Those born between 1955 and 1965 (orange line) and those born between 1965 and 1975 (blue line). The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Own lifetime income Y_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The number of children includes all biological offspring born before age 45. For more details see Section 3.

Figure A.8: Income Rank of the Spouse



Note: This figure plots the coefficients of our main specification when regressing income rank of individual i's spouse on their own lifetime income C_i, parental (father + mother) lifetime income B_i, and the interaction in regression XXX. The estimation sample is 988,302 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The estimation sample is smaller than the main sample as we condition on having a spouse at the age of 45. We can identify spouses if they are married or cohabitating with a joint child. We define the income rank of the spouse as the sum of all labor income between ages 25 and 45, ranked in the population of Sweden. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual.

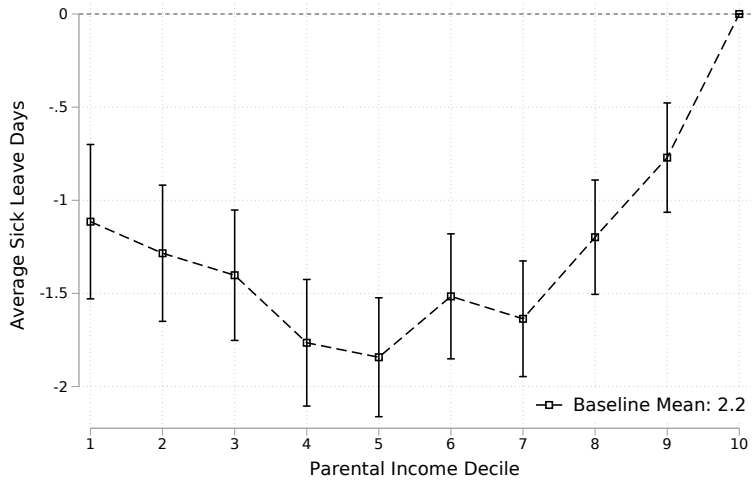
Figure A.9: Mortality



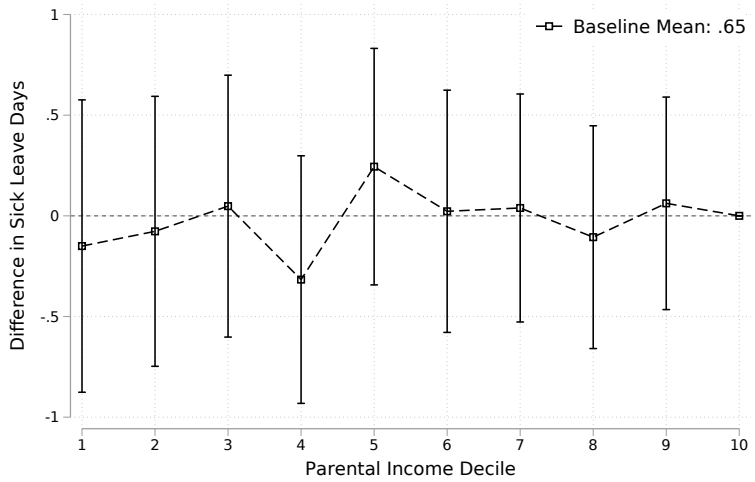
(a) Interaction, β_l

Note: **Panel A.9a** illustrates the relative difference in mortality between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and their parental (father + mother) lifetime income B_i in regression XXX where we interact the control variables by gender g. The outcome variable K_i, mortality, is defined as a indicator function that takes the value one if the individual dies at an age before the last observed age of the respective birth cohort. For example, for the cohort born in the year 1970 this would imply that the individuals dies at an age younger than age 50 which is the last age we observe the 1970-cohort in our data at year 2020. **Panel ??** and **Panel ??** display the corresponding estimates of own lifetime income, $\{\delta_l\}$, and parental lifetime income, $\{\gamma_l\}$, respectfully. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. We report the mean for individuals in the top-decile income bracket with parents in the top-decile income bracket. For more details see Section 3.

Figure A.10: Days of Sick Leave



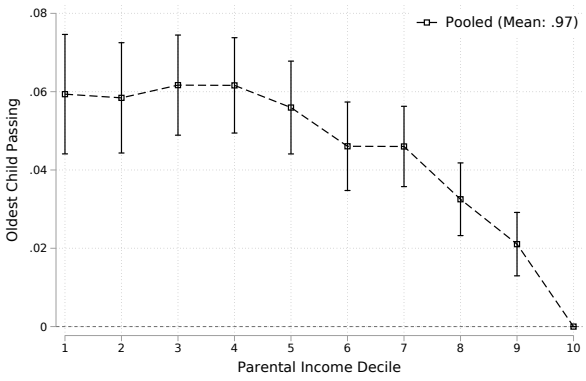
(a) Level



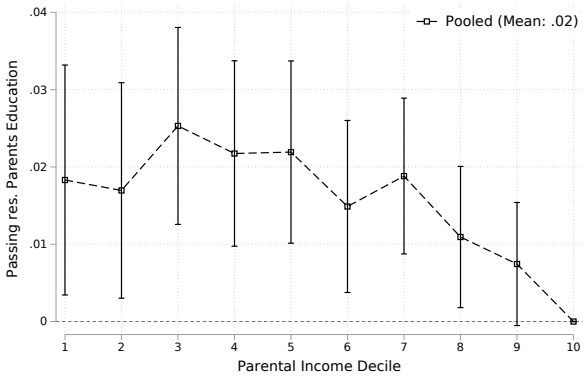
(b) Change

Note: **Panel A.10b** illustrates the relative difference in sick leave days between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and their parental (father + mother) lifetime income B_i, where we interact the control variables by gender g. The outcome variable K_i, difference in sick leave days, is defined as the change in the average number of days on sick leave per year measured between the ages of 40 and 45 and the the ages of 35 and 40. We exclude sick leaves related to pregnancy. In these regressions, we additionally control for the year of birth and the month of birth. **Panel ??** and **Panel ??** display the corresponding estimates of own lifetime income, $\{\delta_l\}$, and parental lifetime income, $\{\gamma_l\}$, respectfully. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Due to data limitations, we can measure the days of sick leave for the cohorts born in 1959 to 1975. We report the mean for individuals in the top-decile income bracket with parents in the top-decile income bracket. For more details see Section 3.

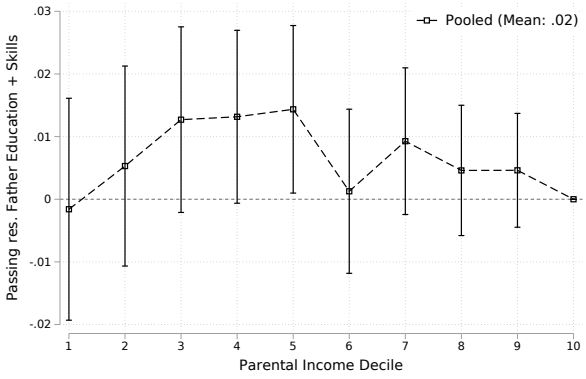
Figure A.11: Children's Educational Achievements: Passing All Subjects



(a) Children's Education



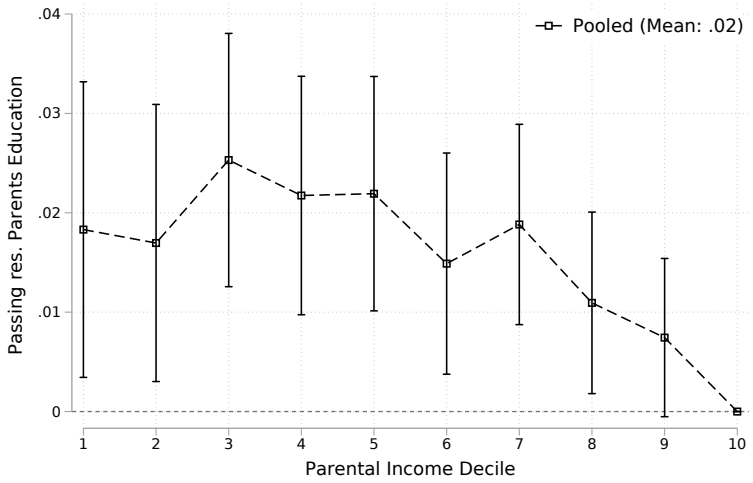
(b) Children's Education, controlling for parental education



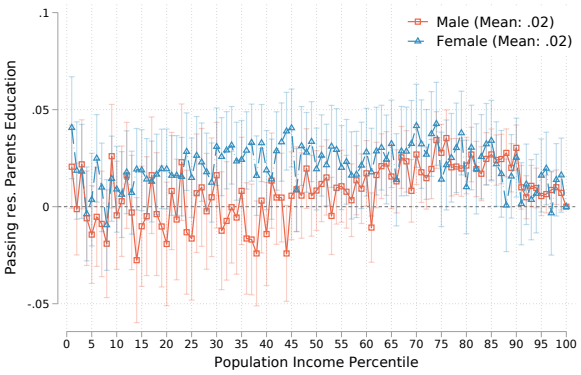
(c) Children's Education, controlling for parental education & skills

Note: **Panel ??** illustrates the relative difference in school performance of the oldest child between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We residualize the indicator variable by first regressing the child's school performance on both the father's and mother's education level and field of study in panel b. We residualize the indicator variable by first regressing the child's school performance on the father's skills including the education level, the field of study, and measures for cognitive and non-cognitive skills in panel c. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and their parental (father + mother) lifetime income B_i where we interact the control variables by gender g. The outcome variable K_i, the school performance of the oldest child, is defined as an indicator that takes the value one if the individual's oldest child passes all three subjects in the national exam in grade 9, that is the child passes the English, Swedish, and Mathematics exams. In the regression, we additionally control for the year of birth and the child's year of birth. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. In this figure, we restrict the sample to individuals that have at least one child and to individuals born between 1963 to 1969 due to data limitations. We report the mean for individuals in the top-decile income bracket with parents in the top-decile income bracket. For more details see Section 3.

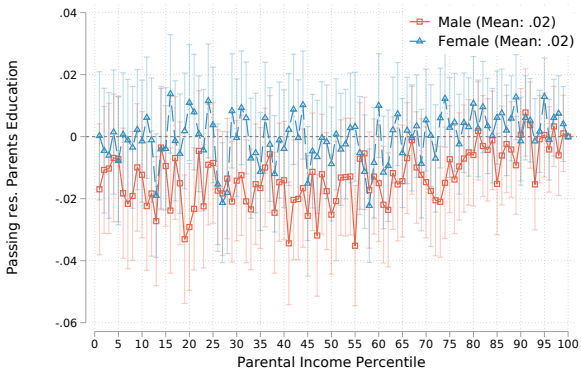
Figure A.12: Children's Educational Achievements Controlling for Parents' Education



(a) Interaction, β_l



(b) Own Income Pct Fixed Effects, δ_l



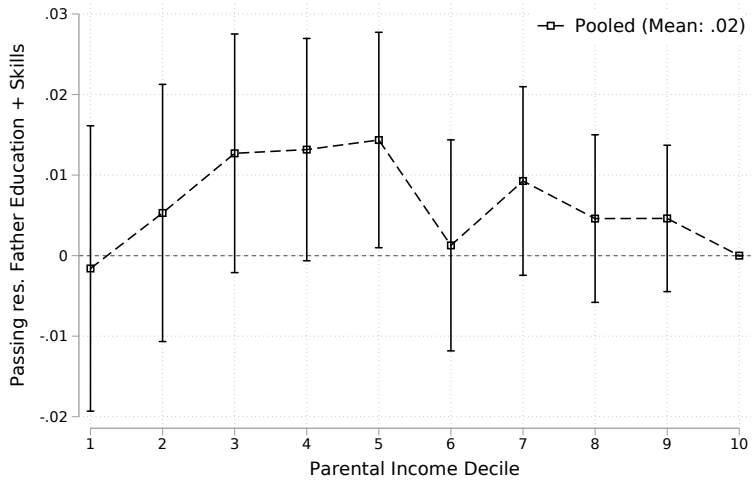
(c) Parental Income Pct Fixed Effects, γ_l

Note: **Panel A.12a** illustrates the relative difference in residualized school performance of the oldest child between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and their parental (father + mother) lifetime income B_i:

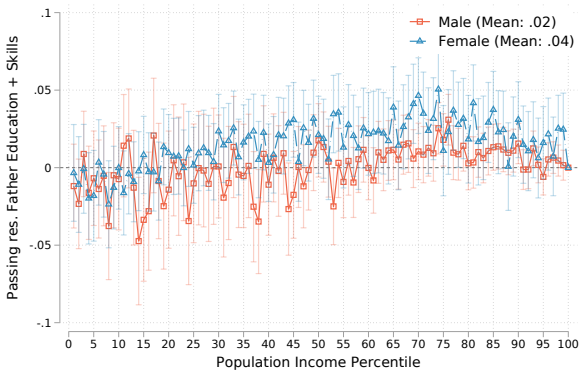
$$K_i = \sum_{g\in\{0,1\}} \left[\alpha^g + \sum_{l=1}^{99} \underbrace{\delta_l^g \mathbf{1}\left(R_{100}\left(C_i\right)=l\right)}_{\text{Own LT Income}} + \underbrace{\gamma_l^g \mathbf{1}\left(R_{100}\left(B_i\right)=l\right)}_{\text{Parental LT Income}} \right] \times \mathbf{1}\left(G_i=g\right) + \sum_{l=1}^{9} \beta_l \underbrace{\mathbf{1}\left(R_{10}\left(C_i\right)=10\right)}_{\substack{\text{Own Income}\\ \text{Top Decile}}} \times \underbrace{\mathbf{1}\left(R_{10}\left(B_i\right)=l\right)}_{\substack{\text{Parental Income}\\ \text{Deciles}}} + \varepsilon_i$$

where we interact the control variables by gender g. We additionally control for the year of birth and the child's year of birth. The outcome variable K_i, the residualized school performance of the oldest child, is defined as a residualized indicator that takes the value one if the individual's oldest child passes all three subjects in the national exam in grade 9, that is the child passes the English, Swedish, and Mathematics exams. We residualize the indicator variable by first regressing the child's school performance on both the father's and mother's education level and field of study in a separate step. **Panel A.12b** and **Panel A.12c** display the corresponding estimates of own lifetime income, $\{\delta_l\}$, and parental lifetime income, $\{\gamma_l\}$, respectfully. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. In this figure, we restrict the sample to individuals that have at least one child and to individuals born between 1963 to 1969 due to data limitations. We report the mean for individuals in the top-decile income bracket with parents in the top-decile income bracket. For more details see Section 3.

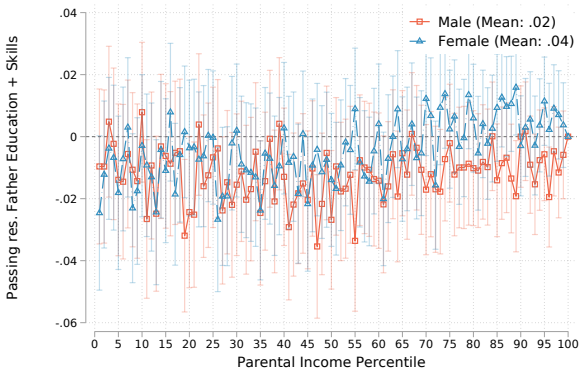
Figure A.13: Children's Educational Achievements Controlling for Father's Skills



(a) Interaction, β_l



(b) Own Income Pct Fixed Effects, δ_l



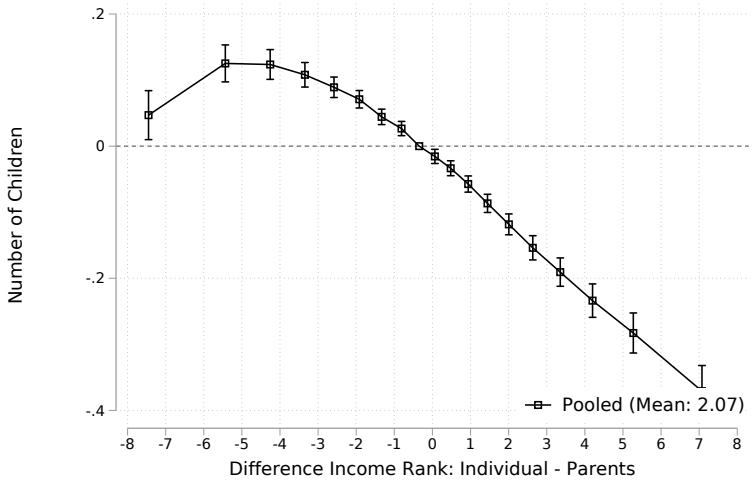
(c) Parental Income Pct Fixed Effects, γ_l

Note: **Panel A.13a** illustrates the relative difference in residualized school performance of the oldest child between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and their parental (father + mother) lifetime income B_i:

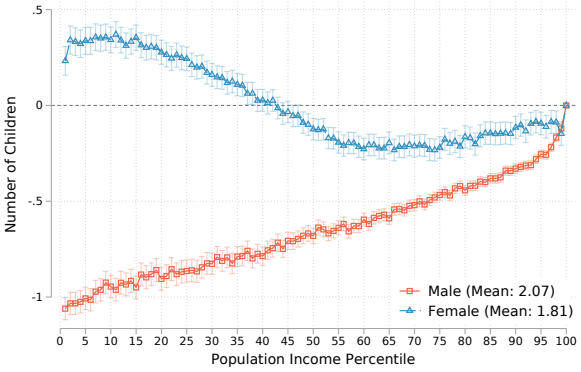
$$K_i = \sum_{g\in\{0,1\}} \left[\alpha^g + \sum_{l=1}^{99} \underbrace{\delta_l^g \mathbf{1}\left(R_{100}\left(C_i\right)=l\right)}_{\text{Own LT Income}} + \underbrace{\gamma_l^g \mathbf{1}\left(R_{100}\left(B_i\right)=l\right)}_{\text{Parental LT Income}} \right] \times \mathbf{1}\left(G_i=g\right) + \sum_{l=1}^{9} \beta_l \underbrace{\mathbf{1}\left(R_{10}\left(C_i\right)=10\right)}_{\substack{\text{Own Income}\\ \text{Top Decile}}} \times \underbrace{\mathbf{1}\left(R_{10}\left(B_i\right)=l\right)}_{\substack{\text{Parental Income}\\ \text{Deciles}}} + \varepsilon_i$$

where we interact the control variables by gender g. We additionally control for the year of birth and the child's year of birth. The outcome variable K_i, the residualized school performance of the oldest child, is defined as a residualized indicator that takes the value one if the individual's oldest child passes all three subjects in the national exam in grade 9, that is the child passes the English, Swedish, and Mathematics exams. We residualize the indicator variable by first regressing the child's school performance on the father's skills including the education level, the field of study, and measures for cognitive and non-cognitive skills in a separate step. **Panel A.13b** and **Panel A.13c** display the corresponding estimates of own lifetime income, $\{\delta_l\}$, and parental lifetime income, $\{\gamma_l\}$, respectfully. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The estimation sample is based on 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. In this figure, we restrict the sample to individuals that have at least one child and to individuals born between 1963 to 1969 due to data limitations. We report the mean for individuals in the top-decile income bracket with parents in the top-decile income bracket. For more details see Section 3.

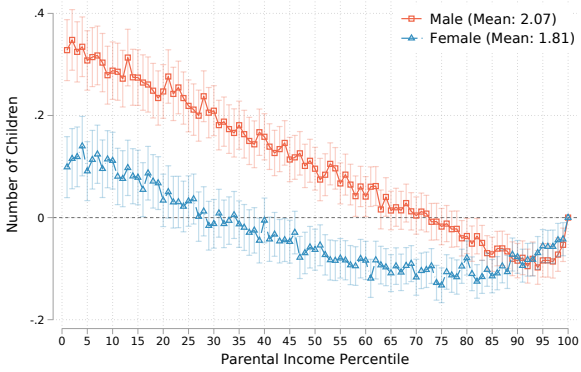
Figure A.14: Cost of Mobility when Rising above Parents' Rank



(a) Binned Difference in Income Ranks, ν_l



(b) Own Income Pct Fixed Effects, δ_l



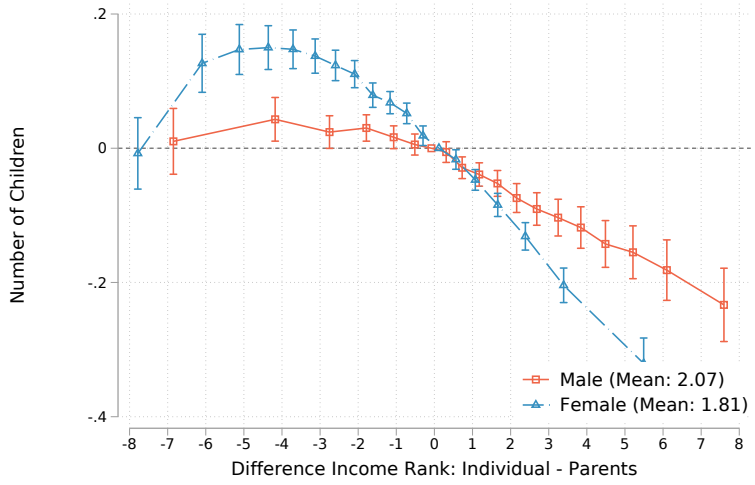
(c) Parental Income Pct Fixed Effects, γ_l

Note: This figure displayed the association between the number of children and the relative income difference between the individual and the parents controlling non-parametrically for own individual lifetime income and parental lifetime income in percentiles. The relative income rank describes the difference in the individual's lifetime income rank and the parents' lifetime income rank. We categorize the relative difference into 19 equally sized bins, and run the following regression:

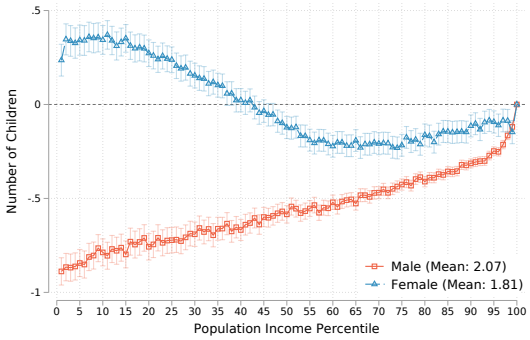
$$\underbrace{K_i}_{\text{Number of Children}} = \sum_{g\in\{0,1\}} \left[\alpha^g + \sum_{l=1}^{99} \underbrace{\delta_l^g \, \mathbf{1}\left(R_{100}\left(C_i\right)=l\right)}_{\text{Own LT Income}} + \underbrace{\gamma_l^g \, \mathbf{1}\left(R_{100}\left(B_i\right)=l\right)}_{\text{Parental LT Income}} \right] \times \mathbf{1}\left(G_i=g\right) + \sum_{l=1}^{19} \nu_l \underbrace{\mathbf{1}\left(R\left(C_i\right)-R\left(B_i\right)=l\right)}_{\text{Binned Rank Diff}} + \varepsilon_i$$

, where we interact the control variables with gender g. We omit the category of having equal ranks as the reference group. The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The number of children includes all biological offspring before the age of 45.

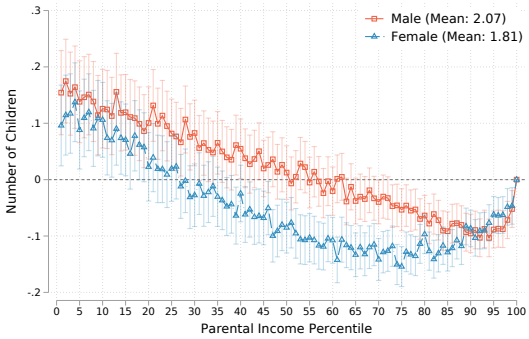
Figure A.15: Cost of Mobility when Rising above Parents' Rank



(a) Binned Difference in Income Ranks, ν_l



(b) Own Income Pct Fixed Effects, δ_l



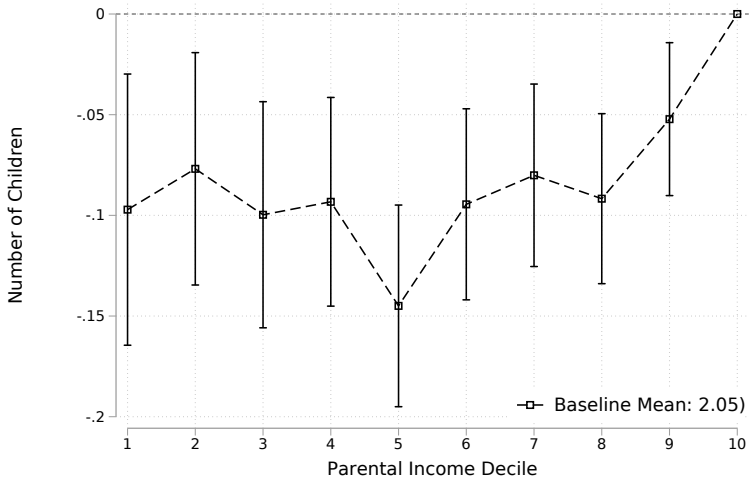
(c) Parental Income Pct Fixed Effects, γ_l

Note: This figure displayed the association between the number of children and the relative income difference between the individual and the parents controlling non-parametrically for own individual lifetime income and parental lifetime income in percentiles. The relative income rank describes the difference in the individual's lifetime income rank and the parents' lifetime income rank. We categorize the relative difference into 19 equally sized bins by gender, and run the following regression separately by gender:

$$\underbrace{K_i}_{\text{Number of Children}} = \alpha + \sum_{l=1}^{99} \underbrace{\delta_l \, \mathbf{1}\left(R_{100}\left(C_i\right) = l\right)}_{\text{Own LT Income}} + \underbrace{\gamma_l \, \mathbf{1}\left(R_{100}\left(B_i\right) = l\right)}_{\text{Parental LT Income}} + \sum_{l=1}^{19} \nu_l \underbrace{\mathbf{1}\left(R\left(C_i\right) - R\left(B_i\right) = l\right)}_{\text{Binned Rank Diff}} + \varepsilon_i$$

, where we omit the category of having equal ranks as the reference group. The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. Own lifetime income Y_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The number of children includes all biological offspring before the age of 45.

Figure A.16: Within Family Result: mother FEs



(a) Interaction, β_l

Note: **Panel A.16a** illustrates the relative difference in the number of children between individuals reaching the top-decile of lifetime income and individuals in the bottom 90 percent separately by parental income deciles. We use parents in the top income decile as the reference group. Formally, it plots the coefficients of the interaction terms, $\{\beta_l\}$, in the regression of outcome K_i of individual i on their own lifetime income C_i and family fixed effects, where θ^g_{mother} are fixed effects indicating sharing the same biological mother. The estimation sample is 682,486 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. The estimation sample is smaller than the main sample as we only include families with more than one same-sex sibling. This is required in order to identify the mother fixed effect between siblings. We run the regression jointly by gender but fully interacted except for the main coefficient of interest (β_l). Own lifetime income C_i is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income B_i is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. We rank these income measures within the birth cohort of the individual. The number of children includes all biological offspring born before age 45. For more details see Section 3

C Appendix Tables

Table A.1: Sample Characteristics

	Mean	SD	Min	Max
Female	0.48	(0.50)	0.00	1.00
Year of Birth	1965.14	(6.01)	1955	1975
Number of Siblings	1.90	(1.25)	0.00	7.00
Income Population Rank	0.50	(0.29)	0.00	1.00
Parental Income Rank	0.50	(0.29)	0.00	1.00
Number of Children	1.90	(1.12)	0.00	5.00
Any Children	0.85	(0.36)	0.00	1.00
Days of Parental Leave per Child | Having Children	219.74	(171.44)	0.00	1462.50
Age at first Child	28.68	(5.48)	17.00	44.00
Years of Education	12.51	(2.26)	7.00	20.00
Education >= College	0.34	(0.47)	0.00	1.00
Age at first Employment	24.82	(4.69)	20.00	60.00
AKM Firm Fixed Effect at first Job	0.10	(0.32)	-5.75	6.63
Number of Firms btw. age 30-45	3.03	(1.79)	1.00	15.00
Observations	1590709			

Note: The table reports sample characteristics of the main estimation sample. The estimation sample is 1,590,709 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden.

Table A.2: Health Outcomes

	(1) Dying before End of Period	(2) Days of total sick leave
Own Top Decile x Parents Decile 1-9	-0.0015	-18.6392***
	(0.0009)	(1.7133)
Individual Income Rank x Female	X	X
Parents Income Rank x Female	X	X
YOB x MOB x Female	X	X
Dep. Mean (Top 10% x Parents top 10%)	.01	19.298
N	1,590,709	1,590,709
R-Squared	0.0213	0.1177

Standard errors in parentheses. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Note: This table shows the regression output of running a modified specification of Equation 1 where we bundle the interaction term and pool by gender:

$$K_i = \sum_{g \in \{0,1\}} \left[\sum_{l=1}^{99} \underbrace{\delta_l^g \mathbf{1}\left(R_{100}\left(C_i\right)=l\right)}_{\text{Own LT Income}} + \underbrace{\gamma_l^g \mathbf{1}\left(R_{100}\left(B_i\right)=l\right)}_{\text{Parental LT Income}} \right] \times \mathbf{1}\left(G_i = g\right) + \beta \underbrace{\mathbf{1}\left(R_{10}\left(C_i\right)=10\right)}_{\substack{\text{Own Income} \\ \text{Top Decile}}} \times \underbrace{\mathbf{1}\left(R_{10}\left(B_i\right) \in \{1,..,9\}\right)}_{\substack{\text{Parental Income} \\ \text{Bottom 90}}} + \varepsilon_i$$

where K_i captures mortality (Column (1)) or days of total sick leave (Column (2)). Mortality is defined as an indicator function that takes the value one if the individual dies at an age before the last observed age of the respective birth cohort. For example, for the cohort born in the year 1970, this would imply that the individuals dies at an age younger than age 50 which is the last age we observe the 1970-cohort in our data in the year 2020. Days of sick leave is defined as the average number of days on sick leave per year between the ages of 35 and 45. We exclude sick leaves related to pregnancy. Days of total sick leave include all sick leaves from work that take longer than two weeks.

Table A.3: SUN2000 - 4 digit education field codes

Code	**Description (English Translation)**
010a	Broad, general education, general orientation
010b	Broad, general education, social science and humanities focus
010c	Broad, general education, natural science focus
010x	Broad, general education, other specialization
080z	Learning to read and write for adults
090z	Personal development
140z	Pedagogy and teacher training, general education
142z	Pedagogy
143a	Teacher training, specialization preschool
143b	Teacher education, focus on leisure activities
143x	Other teacher training for preschool and leisure activities
144a	Teacher training, primary school lower ages, general subject orientation
144b	Teacher training, primary school lower ages, mathematics and nature-oriented subjects
144c	Teacher training, primary school lower ages, Swedish and community-oriented subjects
144x	Other teacher training for the lower ages of primary school
145a	Subject teacher training, mathematics, data and science
145b	Subject teacher training, history, social science and economics
145c	Subject teacher training, language
145d	Subject teacher education, mathematics, data, and natural sciences in combination with history, social sciences and economics
145e	Subject teacher training, mathematics, data, and natural sciences in combination with languages
145 f	Subject teacher training, history, social science and economics in combination with languages
145g	Subject teacher training, mathematics, data and science in combination practical-aesthetic subjects
145h	Subject teacher training, history, social science and economics in combination with practical-aesthetic subjects
145i	Subject teacher training, language, in combination with practical-aesthetic subjects
145x	Other subject teacher training
146a	Teacher training, industry and crafts
146b	Teacher training, trade and administration
146c	Teacher training, hotel and restaurant
146d	Teacher training, food and confectionery
146th	Teacher training, natural farming

146 f	Teacher training, care and social care
146g	Teacher training, media production
146h	Teacher education, music
146i	Teacher training, visual and visual arts
146j	Teacher training, dance and theatre
146k	Teacher training, wood and metal work and textile work
146l	Teacher training, sports and health
146m	Teacher training, household science
146x	Teacher training in other vocational subjects and practical-aesthetic subjects
149a	Special education teacher training
149b	Home language and mother tongue teacher training
149c	School leader training
149x	Other education in pedagogy and teacher training
210z	Art and media, general education
211a	Visual and visual arts education
211b	Art science and art history
211x	Other education in the visual and visual arts
212a	Music
212b	Dance, theater and drama
212c	Music, dance, film and theater science and history
212x	Other education in music, dance and drama
213a	Media production, general education
213b	Film, radio and TV production
213c	Graphic technology and bookbinding
213d	Illustration, advertising, graphic design and photography
213x	Other education in media production
214a	Fashion design
214b	Interior design
214c	Industrial and product design
214x	Other education in design
215z	Crafts
219z	Art and media, other and unspecified education
220z	Humanities, general education
221a	Theology bachelor's degree and equivalent
221b	Pastor, missionary and deacon training
221x	Other education in religion
222a	General language skills and general language education
222b	English
222c	German
222d	French
222e	Spanish
222 f	Russian
222g	Swedish as a foreign language
222x	Other education in foreign languages
223a	Swedish and Nordic languages

223b	Other mother tongues and sign languages
223c	Literary studies and literary history
223x	Other education in Swedish and literary studies
225a	History
225b	Archaeology
225c	Cultural preservation, cultural administration and museum science
225x	Other education in history and archaeology
226z	Philosophy and logic
229z	Humanities, other and unspecified education
310a	Behavioral Science, General Education
310b	Social science, general education
310x	Other broad education in social and behavioral sciences
311a	Psychology education
311b	Psychology training, including completed internship (PTP)
311c	Psychotherapist training
311x	Other education in psychology
312a	Sociology
312b	Social anthropology and ethnology
312c	Cultural and social geography
312x	Other education in sociology, ethnology and cultural geography
313z	Political science
314z	Economics and economic history
319z	Social and behavioral sciences, other and unspecified education
320z	Journalism and information, general education
321a	Journalism
321b	Media and communication science
321x	Other education in journalism and media science
322a	Librarian education and library and information science
322b	Documentation and information science
322x	Other education in library and documentation science
329z	Journalism and information, other and unspecified education
340a	Economics education and business economics
340b	Business and Administration
340x	Other broad education in business administration, trade and administration
341z	Purchasing, sales and distribution
342z	Marketing
343z	Banking, insurance and financing
344z	Accounting and Taxation
345a	Leadership, organization and governance
345b	Administration and management
345c	Personnel management
345x	Other training in management and administration
346a	Medical secretarial training
346x	Other training in office services and secretarial services
347z	Workplace-related training

349z	Business administration, trade, administration, other and unspecified education
380a	Legal education (Bachelor of Laws and equivalent)
380b	Legal education including completed legal qualification
380x	Other education in law and jurisprudence
420z	Biology and environmental science, general education
421a	Biology
421b	Biochemistry, toxicology, pharmacology and nutrition
421x	Other education in biology and biochemistry
422z	Environmental science
429z	Biology and environmental science, other and unspecified education
440z	Physics, chemistry and earth science, general education
441z	Physics
442z	Chemistry
443z	Earth science and natural geography
449z	Physics, chemistry and geoscience, other and unspecified education
460z	Mathematics and science, general education
461z	Mathematics
462z	Statistics
469z	Mathematics and science, other and unspecified education
480z	Data, general education
481a	Systems science and software engineering
481b	Computer science and computer science
481x	Other education in computer science and systems science
482z	Computing
489z	Data, other and unspecified education
520a	Civil engineering education, general technology and technical physics
520b	Engineering education, general technology and technical physics
520x	Other broad education in technology and technical industry
521a	Civil engineering education, majoring in mechanical engineering
521b	Engineering education, majoring in mechanical engineering
521c	Industry and workshop technology
521x	Other education in mechanical engineering and workshop technology
522a	Master's degree in engineering, majoring in energy and electrical engineering
522b	Engineering education, specialization energy and electrical engineering
522c	Electrical installation and electrical mechanics
522d	Energy, operation and maintenance
522x	Other education in energy and electrical engineering
523a	Master's degree in engineering, majoring in electronics, computer technology and automation
523b	Engineering education, majoring in electronics, computer technology and automation
523c	Electronics, telecommunications and computer engineering work
523d	Automation, control and regulation technology
523x	Other education in electronics, computer technology and automation
524a	Civil engineering education, specialization chemical and biotechnology
524b	Engineering education, specialization chemical and biotechnology
524c	Laboratory technology

524d	Process industry technology
524x	Other education in chemical and biotechnology
525a	Master's degree in engineering, specialization vehicle and vehicle technology
525b	Engineering education, focus on vehicle and craft technology
525c	Vehicle technology
525d	Naval and aeronautical engineering
525x	Other education in vehicle and craft technology
526a	Civil engineering education, majoring in industrial economics and organization
526b	Engineering education, majoring in industrial economics and organization
526x	Other education in industrial economics and organization
529a	Civil engineering education, other and unspecified education
529b	Engineering education, other and unspecified education
529x	Other education in technology and technical industry
540a	Master's degree in engineering, majoring in materials technology
540b	Engineering education, specialization materials technology
540x	Other broad training in materials and manufacturing
541a	Master's degree in engineering, food specialization
541b	Engineering education, focus on food
541c	Food manufacturing and handling, general education
541d	Meat and charcuterie
541e	Bakery and patisserie
541x	Other training in food manufacturing and handling
542a	Civil engineering education, focus on textile technology
542b	Engineering education, focus on textile technology
542c	Manufacture of textiles and clothing
542d	Shoe and leather work
542x	Other education in textiles, clothing and leather
543a	Civil engineering education, specialization wood, paper, glass, porcelain and plastic
543b	Engineering education, focus on wood, paper, glass, porcelain and plastic
543c	Wood technical work
543d	Pulp and paper manufacturing
543e	Glass and porcelain manufacturing
543 f	Manufacture of plastic and rubber products
543x	Other training in manufacturing in specific materials
544a	Civil engineering education, focus on rock and mineral technology
544b	Engineering education, focus on rock and mineral technology
544c	Mining and iron mill work
544x	Other education in rock and mineral engineering
549z	Materials and manufacturing, other and unspecified training
580a	Civil engineering education, broad focus on construction, i.a. land surveying
580b	Engineering education, broad focus on construction, i.a. land surveying
580x	Other broad education in civil engineering and building technology
581a	Master's degree in civil engineering, majoring in civil engineering and social environment
581b	Engineering education, focus on civil engineering and social environment
581c	Architectural education

581d	Landscape architect and landscape engineer education
581x	Other education in civil engineering and architecture
582a	Civil engineering education, majoring in construction technology
582b	Engineering education, majoring in construction technology
582c	Construction work, general education
582d	Building carpentry
582e	Painting
582 f	Plumbing training
582g	Concrete, construction and road work
582 h	Masonry education
582 i	Sheet metal warehouse
582x	Other education in building and civil engineering
589z	Civil engineering and building technology, other and unspecified education
620z	Agriculture, gardening, forestry and fishing, general education
621a	Agronomist training
621b	Landmaster training
621c	Education in agriculture, general
621d	Education in agriculture and horticulture
621e	Education in animal husbandry
621 f	Training in horse husbandry and hippologist training
621g	Education in animal care
621x	Other education in agriculture
622a	Horticulture training
622b	Horticultural engineering and horticultural technician training
622c	Training in horticulture
622x	Other education in gardening
623a	Master hunter and forest scientist training
623b	Forester, forest technician and forest engineer training
623c	Education in forestry
623x	Other training in forestry
624z	Fishing and aquaculture
629z	Agriculture, gardening, forestry and fishing, other and unspecified education
640a	Veterinary education
640x	Other training in animal health care
720z	Health care, general education
721a	Medical education
721b	Medical education, including completed public service (AT)
721c	Medical training, specialist competence in surgery, anesthesia and intensive care
721d	Medical education, specialist competence in internal medicine
721e	Medical training, specialist competence in children's medicine
721 f	Medical education, specialist competence in geriatrics
721g	Medical education, specialist competence in general medicine
721 h	Medical training, specialist competence in psychiatry
721 i	Medical education, specialist competence in radiology
721j	Medical education, clinical laboratory specialties

721k	Medical education, other specialist skills
721x	Other education in medicine
723a	Nursing education, basic education
723b	Nursing education, general health care
723c	Nursing training, anaesthesia, intensive care, emergency, accidents, surgery
723d	Nursing education, children and young people
723e	Nursing education, geriatric care
723 f	Nursing education, psychiatric care
723g	Nursing education, district nursing activities
723 h	Nursing education, midwifery
723i	Nursing education, diagnostic radiology and oncology
723j	Nursing education, other specialization
723m	Other nursing education, general care and social care
723n	Other nursing education, focusing on healthcare
723 o	Other nursing education, focusing on psychiatric care
723p	Other nursing education, focusing on elderly care
723q	Other nursing education, focusing on the mentally retarded
723r	Other nursing education, focusing on the disabled
723p	Other nursing education, other orientations
723x	Other education in nursing
724a	Dental education
724b	Dental education, including specialization
724c	Dental hygienist training
724d	Dental nursing education
724e	Dental technician training
724x	Other education in dentistry
725a	Optician training
725b	Audiology training
725c	Orthopedic engineering education
725d	Biomedical analyst training
725e	Hospital physicist training
725 f	X-ray nursing training
725x	Other technically oriented nursing education
726a	Physiotherapy education
726b	Occupational therapist training
726c	Speech therapy training
726d	Dietician training
726x	Other training in therapy, rehabilitation and dietary treatment
727a	Pharmacist education
727b	Recipe training
727x	Other education in pharmacy
729z	Health care, other and unspecified education
760z	Social work and care, general education
761a	Children and youth, general education
761b	Child care

761c	Leisure activities for children and young people
761x	Other education in children and youth
762a	Social service and care, general education
762b	Social care education, elderly and disabled
762c	Social care training, social pedagogical treatment
762d	Sociology training
762e	Sociology education, elderly and disabled
762 f	Sociologist training, social pedagogical treatment
762g	Study and vocational guidance training
762x	Other education in social work and guidance
769z	Social work and care, other and unspecified education
810z	Personal services, general education
811a	Hotels, restaurants and large households, general education
811b	Training for hotels
811c	Training for restaurant
811d	Training for large households
811x	Other training in hotels, restaurants and large households
812z	Tourism, travel and leisure
813z	Sports and wellness
814a	Home and Household Science, General Education
814b	Training for housekeeping
814x	Other training in household services and housekeeping
815a	Training for hair care
815x	Other training in hair and beauty care
819z	Personal services, other and unspecified training
840a	Transport services, general education
840b	Training for land transport
840c	Training for shipping
840d	Training for aviation
840e	Training for postal service
840x	Other training in transport services
850z	Environmental protection and environmental protection, general education
851z	Environmental technology and environmental control
852z	Nature conservation and animal welfare
853z	Cleaning and waste management
859z	Environmental care and environmental protection, other and unspecified education
860z	Security services, general education
861a	Police training
861b	Training for fire protection and other emergency services
861c	Civil defense training
861x	Other training in community safety
862z	Work environment and worker protection
863a	Professional officer training
863b	Reserve officer training
863c	Tactical training for military personnel

863d	Vocational training for military personnel
863e	Staff training for military personnel
863 f	Higher management training for military personnel
863x	Other military training
869z	Security services, other and unspecified training
999z	Unknown

Note: The table reports the descriptions of the 4 digit education field of study codes (336 categories).

Table A.4: SUN2000 - 3 digit education level codes

Code	Description (English Translation)
000	Other and unspecified preschool education
001	Preschool
002	Preschool class
100	Other and unspecified upper secondary education shorter than 9 years
102	Primary school education, grades 1-6
106	Public school education
200	Other and unspecified pre-secondary education, 9 (10) years
204	Secondary school education
206	Primary school education, grades 7-9
310	Other and unspecified upper secondary education, shorter than two years
312	Secondary education shorter than two years, theoretical and study preparation, no final grade
313	Secondary education shorter than two years, vocationally oriented, no final grade
316	Secondary education shorter than two years, theoretical and study preparation
317	Secondary education shorter than two years, vocationally oriented
320	Other and unspecified upper secondary education, two years
322	Secondary education two years, theoretical and study preparation, no final grade
323	Secondary education two years, vocational, no final grade
326	Secondary education two years, theoretical and study preparation
327	Secondary education two years, vocational
330	Other and unspecified upper secondary education, three years
332	High school education three years, theoretical and study preparation, no final grade
333	High school education three years, vocational, no final grade
336	High school education three years, theoretical and study preparation
337	High school education three years, vocational
410	Other and unspecified post-secondary education, shorter than two years
412	Higher education, 20 credits or 30 higher education credits, not a degree
413	Upper secondary education
415	Post-secondary education shorter than two years, not university/college
417	University education shorter than two years, vocationally oriented
520	Other and unspecified post-secondary education, two years
522	Higher education, 80 credits or 120 higher education credits, not a degree
525	Post-secondary education two years, not university/college
526	University education two years, general
527	University education two years, vocational
530	Other and unspecified post-secondary education, three years
532	Higher education, 120 credits or 180 higher education credits, not a degree
535	Post-secondary education three years, not university/college
536	University education three years, general
537	University education three years, vocational
540	Other and unspecified post-secondary education, four years
545	Post-secondary education four years, not university/college
546	University education four years, general
547	University education four years, vocational
550	Other and unspecified post-secondary education, five years or longer
555	Post-secondary education, five years or longer, not university/college
556	University education five years or longer, general
557	University education five years or longer, vocationally oriented
600	Other and unspecified postgraduate education
620	Licentiate education
640	Doctoral education

Note: The table reports the descriptions of the 3 digit education level codes (47 categories).