Who Can Have It All?

The Unequal Cost of Careers, Inter-generational Mobility, & Equality of Opportunity

Patrizia Massner Arash Nekoei *

April 17, 2025

Abstract

The standard measures of intergenerational mobility rely on occupation or income but overlook the costs of upward mobility. One such cost arises from the career-family tradeoff. We explore the implications of this omission in four parts. Part 1 argues theoretically that unequal career costs (i.e., a steeper tradeoff for disadvantaged individuals) imply that upward mobility in professional life contributes less to welfare mobility if it is achieved at the expense of personal life: income mobility overestimates welfare mobility. Part 2, using Swedish data, shows that, indeed, individuals from low-income families sacrifice more of their family life to attain a career (higher lifetime income). They delay childbearing, have fewer children, and spend less time with each child. Part 3 explains that the unequal cost of career stems from a more arduous path to success, shaped by disadvantages in time and space. To build a career, disadvantaged individuals study longer but in less remunerative fields—despite their stronger cognitive and non-cognitive skills—which delays career and family formation. They are also more likely to leave their hometowns and live farther from their parents. Part 4 links mobility and equality of opportunity (EO). Mobility and EO of outcome diverge on two fronts: Only EO conditions on effort, and EO focuses on circumstances, while mobility centers on hereditary traits. Once the focus shifts to welfare, only the latter distinction remains. Thus, whether welfare mobility is a reliable proxy for EO hinges on a normative question: how much hereditary and circumstantial traits overlap.

*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, Matt Notowidigdo, and seminar participants at Nordic Public Finance Conference, IIES, Stockholm University, Helsinki GSE Colloquium, ... for their comments.

Introduction

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

In the social sciences broadly—and in economics in particular—we have extensively studied those who make it from the bottom to the top of the social ladder. We've often interpreted that success, joyfully, as a sign of inclusion. But we have paid far less attention—or perhaps shown far less concern—for how hard people must try, and the costs they bear, to climb that ladder.

In this context, 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, 2014, 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 in four parts. We start with a theoretical model of intergenerational mobility in Part 1. 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 personal life.

Part 2 empirically investigates the unequal cost of a career using Swedish register data. As a proxy for the personal life, we focus primarily on family outcomes. Our main measures of

[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.

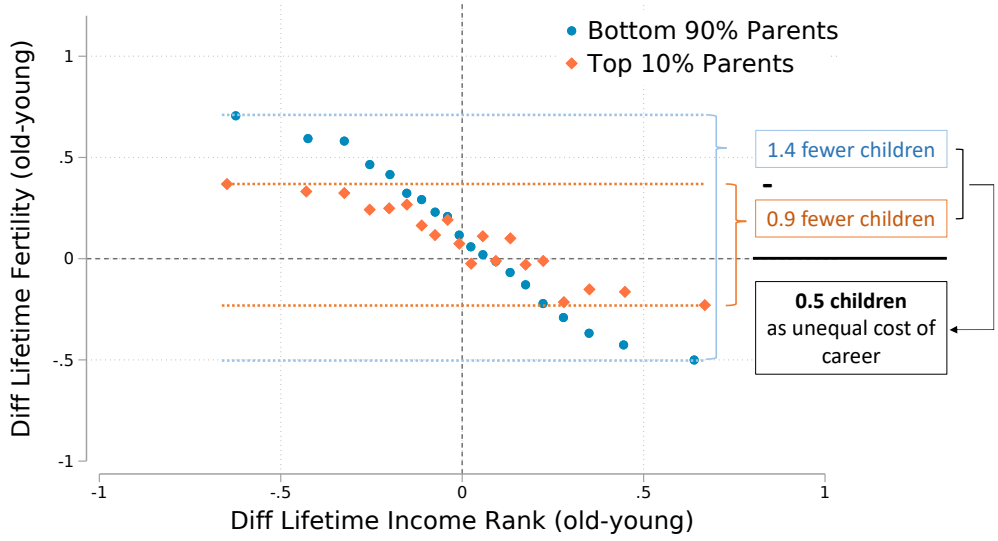
[2] See Corak (2006); Jäntti et al. (2006); Björklund et al. (2009); Blanden (2013); Chetty (2021).

[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 and Ferrie (2013); Berger et al. (2023).

family outcomes are the completed (lifetime) fertility rate and the time spent on parental leave per child. We complement this analysis with other aspects of personal life, such as geographical mobility. Professional life and parental background are measured through an individual's and their parents' lifetime income, similar to the mobility literature (Chetty et al., 2014).

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 outcomes of sisters by plotting the difference in their lifetime fertility (completed fertility rate) against the difference in their lifetime income rank. The former is measured as all biological offspring born before the age of 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 difference 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 .1 fewer children (corresponding to 5 percent fewer children). The extensive and intensive margins add up to .15 fewer children (7.5 percent).

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.

Part 3 shows that the unequal cost of career stems from contrasting trajectories to success, shaped by drawbacks in time and space. We show this with the help of our diff-in-diff approach and detailed data on education attainment, cognitive and non-cognitive skills from the military, labor market outcomes, and geographical location.

To reach the top income decile, individuals with low-income parents postpone childbirth. We find that they are, on average, two years older when giving birth for the first time. This difference entirely explain the difference in the number of children along the intensive margin: Controlling for the age of parenthood, self-made achievers have the same number of children as achievers coming from high-income families.

Postponement of childbirth is associated with compounding disadvantages over time. We find that individuals from low-income backgrounds who pursue a career obtain, on average, one additional year of education compared to their counterparts from high-income families. Furthermore, they delay labor market entry, as indicated by a higher age at first job.

In addition to drawbacks in time, they also have drawbacks across space. Self-made achievers are more likely to leave their birthplace and move geographically more often in their lives. They end up living further from their hometowns and parents compared to achievers from privileged backgrounds. This can be seen as another unequal cost of career, where individual choice of career implies less help from grandparents in growing the kids, which elevates the cost of parenthood.

These drawbacks in time and space materialize into different educational and professional paths to success between individuals from high-income vs. low-income backgrounds. Disadvantaged achievers 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.

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.

Finally, Part 4 turns to how intergenerational mobility (IGM) of welfare can connect the two concepts of intergenerational mobility with equality of opportunity (EO). We begin by demonstrating that the two concepts diverge along two axes. First, EO conditions on effort, prompting much of its literature to grapple with the arduous task of measuring effort (Roemer and Trannoy, 2015). Second, EO addresses unequal outcomes due to circumstances—innate traits deemed morally arbitrary and beyond individual control—whereas mobility focuses on hereditary traits. The stronger the role of heredity, the lower the mobility.

We argue that the first distinction fades when the focus shifts from outcomes to welfare. In a welfare-based framework, effort is endogenous and thus irrelevant as a conditioning variable. EO of outcome is equal outcome for equal effort, EO of welfare is equal outcome for equal cost. What remains is the divergence in the trait sets each approach considers. This insight yields two lessons.

First, both concepts concern hereditary circumstances, like race, family ties, or last name. On one hand, only EO focuses on non-hereditary circumstances, such as gender. This explains that gender discrimination is part of EO but not mobility. On the other hand, only IGM accounts for hereditary traits that do not qualify as circumstances. For instance, if traits like height or IQ are rewarded and such remuneration is deemed morally acceptable, EO permits it—yet mobility still reflects it. The more the market rewards such traits, the lower the mobility. This is a shortcoming of mobility, an unintended consequence of the simplification that mobility imposes to facilitate measurement. This simplification also reveals why higher mobility is not invariably a socially desirable outcome.

Second, it implies that the two concepts coincide in a world where almost all hereditary traits coincide with circumstances. However, if circumstances form a set distinct from hereditary traits, the two concepts become independent—implying that a society with higher mobility does not necessarily exhibit greater equality of opportunity.

While identifying hereditary traits is a positive question for scientists, defining circumstances is a normative, philosophical, and political one. These two axes delineate the conceptual gap between EO and intergenerational mobility.

The paper unfolds as follows. Section 1 delves into the cost of career from a theoretical perspective. Section 2 does the same from an empirical perspective after introducing the data and empirical design. Section 3 investigates the mechanisms behind the unequal cost of a career. Section 4 shows how our findings can connect IGM and EO and investigate when IGM is a reliable surrogate for EO. Section 5 concludes the paper. We relegate additional background material and auxiliary analyses to the Online Appendix.

1 Theory

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

We consider an individual with preferences over personal life, k, and professional life, c. Both aspects of life use a L dimensional vector as input—such as time, effort, or mental capacity—to their production functions: $c = C(x)$ and $k = K(y)$. We assume both production functions are increasing and concave. Nature limits the total input available

$$x + y \leq 1,$$

where 1 is a normalization. This resource constraint creates a career-family tradeoff. This model nests 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 $C(x) = wx$ and $K(y) = y$.

The individual lifetime optimization then can be written as:

$$\max_{x,y \in \mathbb{R}^L} U(c, k) \tag{1}$$

where $c = C(x)$ and $k = K(y)$ and under the L constraints: $x + y \leq 1$.

Parental background matters in three ways. First, it can affect the career production function. For example, social connections or the last name can affect the return to effort in the labor market. Second, it can affect the family production function. Third, it can help to relax the resource constraints.

Similarly to converting firms' profit maximization into cost and scale optimization in production theory, we reformulate the individual problem into two-step optimization. In the inner problem, the individual maximizes her personal life given the amount of career that she has to achieve

$$\kappa(c) = \max_{x,y} K(y) \tag{2}$$

under professional life production function $C(x) \geq c$ and the resource constraint $x + y \leq 1$. So, the function $\kappa(c)$ is the optimal personal life given a minimum level of professional life (career cost function, for short). It is decreasing, $\frac{\partial \kappa}{\partial c} < 0$, where the derivative captures the loss of personal life for one unit of career gain.

The outer problem is balancing the trade-off between professional and personal lives

$$\max_{c} U(c, \kappa(c)) \tag{3}$$

The gain from the two-step formulation of individual optimization is that the parental background influence is entirely captured in the career cost function κ.

The key assumption is that the parental career is a sufficient statistic for background (Atkinson, 1980b). We will denote by $\tilde{}$ variables of parent generation. Thus, the individual career vs. family

optimization becomes:

$$\max_{c} U\left(c, \kappa\left(c, \widetilde{c}\right)\right) \tag{4}$$

We can now define the unequal cost of a career as the sub-modularity of the career cost function κ. This is that the individual's and the parent's career are complements:

$$\frac{\partial^2 \kappa}{\partial c \partial \widetilde{c}} < 0 \tag{5}$$

Our main goal is to investigate how the unequal cost of a career, i.e., condition 2, creates a wedge between income and welfare mobility.

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 c. 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 both approaches, we assume a steady-state economy. More precisely, all generations are identical with stable marginal and joint distributions across generations.

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

$$U(c,k) = \alpha \ln\left(c - \gamma_c\right) + \overline{\alpha} \ln\left(k - \gamma_k\right).$$

We further assume the optimal personal life function as

$$\kappa\left(c, \widetilde{c}\right) = \left(1 - \frac{c}{w}\right) \widetilde{c}^{\beta}$$

where productivity follows a Markov chain where $(w, \tilde{w})$ is jointly log-normal distributed with correlation ρ. The case of $\beta = 0$ coincides with the labor supply model at wage rate w. If we assume that $\gamma_c = \gamma_k = 0$, i.e., Cobb-Douglas utility function, then we can analytically address our question.

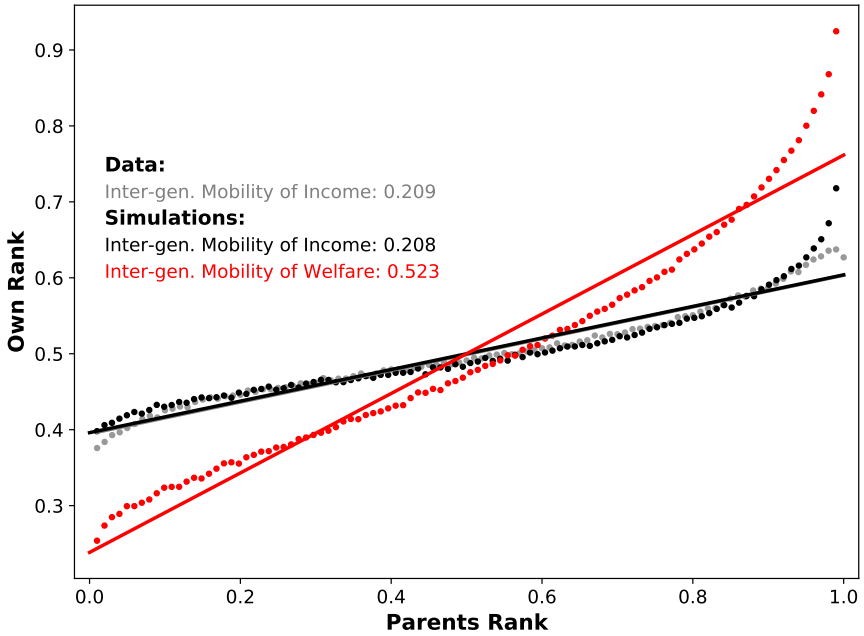
Proposition 1. *In the case where $\gamma_c = \gamma_k = 0$, intergenerational welfare mobility is lower than intergenerational income mobility when there is a biased background $\beta > 0$.*

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 $c = \alpha w$. But the intergenerational mobility in welfare is different as the welfare rank depends on the ranking of a function of own and parents' productivity, more precisely

$$\alpha \ln(w) + \overline{\alpha}\beta \ln\left(\tilde{w}\right).$$

For the case where $\gamma_c, \gamma_k > 0$, 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 model's implied intergenerational correlation of income and welfare. In black (red), we depict the relationship between an individual's own income rank c (welfare rank U) and the parent's income $\tilde{c}$ (welfare rank $\tilde{U}$). For this simulation, we assume Stone-Geary utilities with the normalization $\gamma_c = \gamma_k = 1$. The number of observations in this simulation is 1.5m. We set the intergenerational correlation of income to $\rho = 0.09$ such that it matches the observed intergenerational income correlation using our sample of Swedish register data (depicted in gray). The lines in the graph illustrate the linear fit between an individual's rank and the parent's rank.

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

$$\kappa\left(c, \widetilde{c}\right) \simeq \alpha c + \beta \widetilde{c} - \gamma \widetilde{c} c$$

We assume a symmetric distribution to have zero third central moments and joint normal using Isserlis' theorem. This assures that individuals optimize, leading to the distribution of c and a Markov chain which is stationary.

The intergenerational welfare mobility is lower than the intergenerational income mobility when we have a biased background, meaning $\frac{\partial^2 \kappa}{\partial c \partial \widetilde{c}} < 0$.

2 From Theory to Data

2.1 Empirical Design

We empirically measure the unequal cost of a career, defined in the theoretical section as $\frac{\partial^2 \kappa}{\partial c \partial \widetilde{c}} < 0$, by regressing the lifetime family outcome of an individual, k, on the individual's lifetime income, c, parental lifetime income, $\widetilde{c}$, 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:

$$k = f(c) + g(\tilde{c}) + \sum_{i=1}^{9} \beta_i \underbrace{R(c)}_{\text{Own Income Rank}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{6}$$

where $D(.)$ is decile function and $R(.)$ the rank function. Our coefficients of interest are $\{\beta_i\}$, capturing the interaction terms between individual and parental income. $R(c)$ captures moving up in rank, which we refer to as "having a career". We interact this term with parents' lifetime income deciles, $\mathbb{1}(D(\tilde{c}) = i)$, where we omit the highest decile group as the reference group. Then, the coefficients of interest $\{\beta_i\}$ 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-socioeconomic background sacrifice more of their family life to achieve a career compared to individuals coming from a high-socioeconomic background. Hence, $\{\beta_i\}$ measures the difference in the cost of upward mobility by background.

We complement this specification by running the regression in a fully non-parametric way. To do so, we replace the linear rank in the interaction term with a dummy variable that equals one if the individual's lifetime income is in the top decile:

$$k = f(c) + g(\tilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{7}$$

where $\mathbb{1}(D(c) = 10)$ indicates having a lifetime income in the top decile which we refer to as "having a career". We place particular emphasis on β_1, which captures the difference in outcomes between individuals with a career who come from bottom-decile parents and those from top-decile parents.

To interpret the magnitude of our coefficients, we compare them to the baseline mean of outcomes. Specifically, we compare the size of β_i to the average family outcome of individuals in the bottom decile of both own and parental income.

Finally, we add a series of controls to Regression 7. We flexibly control for family fixed effects, spouses' lifetime income percentiles, and age at first birth. Including father fixed effects mirrors our initial exercise of comparing two sisters in Figure 1.

2.2 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 for-

mer 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. Individual lifetime income rank, c, is measured as the sum of labor income over the ages 25 to 45. We rank incomes 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, $\widetilde{c}$, 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.[4] 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.[5] 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 first birth, age at marriage, 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.[6] This data restriction results in 1,590,709 individuals. In Figure **??**, 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 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, 214 days of parental leave per child.

[4] 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.

[5] Appendix Figure **??** provides further details on the cohort selection.

[6] 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.

2.3 Unequal Cost of Career: Empirical evidence

Figure 4 displays the costs of mobility measured by the lifetime number of children (Completed Fertility Rate) as our main outcome variable. Figures 4a and 4b focus on the total number of children, while Figure 4c separates the results from extensive and intensive margins.

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 similar contribution from each margin. The impact is slightly less pronounced for men than for women, largely attributable to differences in the extensive margin.

The associated costs of mobility increase as individuals move from the bottom to the top of the income distribution, as captured by the linear specification in Figure 4a. Compared to individuals from top-decile backgrounds, women (men) from the bottom-decile parents have, on average, 0.95 (0.41) fewer children.

Underprivileged achievers tend to embark on family life at a significantly later stage. Figure 5 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.1).

The delay in childbearing accounts for the entire disparity in the number of children at the intensive margin. Figure A.2 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 5b and 5c reveals a significant difference in the amount of parental leave per child. Selfmade women take approximately one month less in parental leave per child than those from privileged backgrounds. Furthermore, men are less likely to take any parental leave. However, conditional on taking leave they spend more days on parental leave.

Reduced levels of parental leave 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.

2.4 Additional Results and Discussions

2.4.1 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 7. 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 7 father fixed effects to measure differences in family outcomes between siblings.[7] 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 in a subset of regressions. These controls the income rank percentile dummies of the spouse and age at first birth.

Figure A.2 shows the results where we add father fixed effects in Equation 7. Our main result on the number of children is robust to the inclusion of family fixed effect, despite the magnitude being slightly different. On average, women (men) who have a career have 0.2 (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.

2.4.2 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. This would be in contrast to evidence we saw on parental leave (Figures 5b and 5c). 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

[7] We add father fixed effects to regression rather than substituting the gender-specific controls of parental income for the following reason: Adding father fixed effects restricts our analysis to families with at least one sibling. If we would further interact the father fixed effect with gender we would restrict the analysis to families with same-sex siblings. By adding the father fixed effects to the regression, we can control for family-specifc factors while keeping a similar sample and controls to Equation 7.

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.8 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.8a and A.8b 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 of study, 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 suggests that the positive correlation is driven by the fact that children of selfmade achievers are, on average, more able. Conditional on parental skills, underprivileged achievers do not significantly invest more or less into their offspring.

2.4.3 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 is done in a household. A.5 shows that underprivileged achievers mate with spouses who 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.

2.4.4 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.4). 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.

2.4.5 Private Life: Beyond Family

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.6 and Table A.1 that we do not find a statistically significant correlation on mortality. Second, we estimate Equation 7 using the total days spent on sick leave as an outcome variable. Selfmade achievers have on average 1.5 fewer days of sick leave per year compared to their privileged counterparts (Figure A.7a). 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.7b). 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.

3 Disadvantages and Mechanisms

The unequal cost of a career comes from two contrasting paths to success. In this section, we first document differences in barriers to success, namely time disadvantages and location barriers. We then analyze the mechanisms through which these disadvantages materialize. That is, we compare educational and professional pathways between selfmades and privileged achievers, i.e., achievers from high vs. low socio-economic backgrounds.

3.1 Disadvantages Across Time and Space

Selfmades face barriers to success across time. Figure 5 shows that to reach the top decile of the income distribution, individuals from bottom decile backgrounds study one year longer. This additional year of studies carries over into the timing of the professional and family life. Specifically, we find that disadvantaged achievers start their first full employment on average 0.4 to 0.8 years later and postpone both marriage and childbirth.

Besides drawbacks in time, disadvantaged achievers further experience drawbacks across space. Figure 7 demonstrates that the average number of total moves for underprivileged achievers is 0.08 higher than for their privileged counterparts, corresponding to a 5-8 percent increase in the number of moves during their working life. This goes along with moves that separate selfmade individuals from their families. We measure this twofold: First, Figure 7b 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 bottom-decile parents live on average about 50 km (69 km for females) further away from their parent'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—is discussed in the following section (see Figure 9).

This suggests that individuals from less advantaged backgrounds often need to exhibit greater time and 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.

3.2 Mechanisms

Knowing that disadvantaged individuals face barriers across both time and space, we now examine the mechanisms through which these disadvantages materialize. We proceed systematically. First, we investigate human capital inequalities by asking whether differences in education or skills account for the observed disparities between groups. If not, a residual remains—implying that, conditional on education and skills, the groups experience different returns or outcomes.

Next, we turn to the labor market. We assess whether the groups differ in access to firms and occupations, or whether, within the same firms and occupations, they receive different earnings. Finally, we integrate the human capital and labor market dimensions by examining the residual variation in earnings, conditional on skills, education, firms, and occupations. Naturally, even this final residual may reflect unobserved skills correlated with family background.

Our empirical approach relies on estimating a series of Mincer equations to harmonize the units of the different measures. We then decompose total earnings into components explained by each factor, using the predicted values. We begin with the human capital exercise to illustrate this framework in practice.

3.2.1 Human Capital Inequality: Longer Education, Better Human Capital

To measure the contribution of education and skills, we use four measures of human capital. Utilizing a rich Swedish dataset on education *degree* and *field*, we decompose the differences in the quantity and quality of education. Regarding the degree, we exploit 47 distinct categories of highest completed education, ranging from pre-school level to doctoral degree (Appendix Table A.3). Regarding the field of study, we observe 336 categories, e.g., "Teacher training, industry and crafts" or "Industrial and product design" (Appendix Table A.2).

Moreover, we use the skill measures 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 for women in Appendix Figure A.3.

The challenge of using these 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. 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 7. Figure 8a illustrates these findings.

To reach the top decile, self-made individuals from bottom-decile backgrounds 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.

3.2.2 Labor Market Inequality: The Role of Firms and Occupations

We harmonize labor market measures using the same approach as before. Specifically, we estimate a Mincerian regression of lifetime income on firm and occupation fixed effects. Occupations are categorized into 123 groups, as detailed in Table A.4. To render labor market contributions comparable to those of human capital, we restrict the analysis to men. The male sample includes 42,180 distinct firms.[8] Corresponding estimates for women are reported in Figure A.3b. The regression yields an R-Squared of 0.499.

Figure 8b depicts the role of labor market components for men from disadvantaged backgrounds. To reach the top decile of the income distribution, individuals from the bottom decile select on average more into high-return occupations. This positive occupational contribution to lifetime earnings is partially offset by their concentration in firms that, on average, pay lower returns.

This pattern persists when we control for human capital measures. As shown in Figure 8b, the dominant positive contributor is the combination of lucrative occupations and relatively higher cognitive and non-cognitive skills among disadvantaged achievers, compared to their high-SES counterparts. The largest negative contributor is the regression residual, which likely captures unobserved aspect such as limited access to networks or job search know-how. We elaborate on these mechanisms in the following section.

[8] Industry-level differences is absorbed by the firm fixed effects.

Which occupations facilitate upward mobility for disadvantaged individuals? Table 2 lists the top five overrepresented occupations among achievers from top versus bottom decile parental backgrounds. Notably, the occupational paths differ. High-SES individuals disproportionately enter professions such as medicine and law, whereas low-SES achievers attain high earnings in careers such as becoming a pilot or joining the military.

3.2.3 Steep Job Ladder

After accounting for firms, occupations, and different measures of human capital residual disparities remain between privileged and selfmade achievers. This residual may reflect certain labor market know-how or access to networks that correlates with the background. We address this in Figure 9 which reports different dimensions of the working life. First, Figure 9a documents that selfmades start their career at later ages. To reach the top decile of the income distribution, those with parental income in the lowest decile are, on average, 0.4 to 0.8 years 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.[9] For male selfmade achievers, the first job has a 5 percent lower earnings premium compared with high-earners of highest-decile parents. However, selfmades change jobs more often, 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. This is because the magnitude of the difference in their first firm is close to the magnitude of the growth.

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 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 9b illustrates that self-made achievers are six percentage points less likely to work for a firm any of tehir parents have worked for—a substantial effect, representing nearly 25 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 9c). This pattern supports the notion that privileged achievers can leverage additional resources beyond conventional human capital.

[9] 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.

4 When is Intergenerational Mobility a Reliable Surrogate for 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 (EO) 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.

4.1 Definitions of IGM and EO of Outcomes

We start with a generalized version of our model in Section 1. Individuals decide a vector x, which we refer to as effort, according to

$$V(\sigma) = \max_{x,y \in \mathbb{R}^L} U\left(C(x,\sigma), K(y,\sigma)\right) \tag{8}$$

under the L resource constraints $x + y \leq 1$ given a vector of exogenous parameters (innate traits) $\sigma = \{\sigma_i\}_{i\in\Sigma}$.

To fix ideas, equality of outcome (career) is achieved when all individuals have the same outcome, $\frac{dC}{d\sigma_i} = 0$ for all $i \in I$. Important to note that this was a total differentiation as equality of outcome includes endogenous effort. In contrast, equality of welfare requires $\frac{\partial V}{\partial \sigma_i} = 0$ for all $i \in I$. We focus on the idealized forms of each concept—e.g., perfect equality instead of a level of inequality—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.

IGM of outcome is defined as a zero correlation between the outcomes of two successive generations:

$$Corr(\mathbf{C}(x^*(\widetilde{\sigma}),\widetilde{\sigma}), \mathbf{C}(x^*(\sigma),\sigma)) = 0. \tag{9}$$

This condition is equivalent to requiring equal outcomes across individuals with different hereditary traits:

$$\frac{d\mathbf{C}}{d\sigma_i} = 0 \quad \text{for } i \in \Sigma_H. \tag{10}$$

This formulation relies on a simplification—namely, that all traits are either fully hereditary—such that $\sigma_i = \widetilde{\sigma}_i$ for all $i \in \Sigma_H$—or fully idiosyncratic, that is σ_i is independent of $\widetilde{\sigma}_i$ for all $i \notin \Sigma_H$.[10]

Under EO, inequality in outcome is justifiable, provided that individuals enjoy the "opportunity" for the good or rank in the relevant sense.

EO has been defined in multiple ways across disciplines. One view holds that EO exists

[10] This unrealistic distinction is made for the sake of simplification, but without loss of generality. It stands in contrast to our modeling of innate traits in Section 1.

when, regardless of birth, occupation, or social position, individuals possess in fact—not merely in form—equal chances to fully use their natural endowments of physique, character, and intelligence (Tawney, 1952): "The existence of such opportunities in fact, and not merely in form, depends, not only upon an open road, but upon an equal start" (see page 109).

Another formulation states that "A man ... is free to choose any occupation he is fit for, and when no one else has an advantage over him in that choice by reason of his wealth or social origins, but only by reason of his ability or social deserts" (Plamenatz, 1975, p. 93). A third definition frames EO as eliminating "all the barriers that prevent individuals from obtaining the training necessary to convert the potential talents implicit in their genetic endowments into marketable capacities" (Taubman, 1979, p. 6). A more technical economic interpretation describes EO as a situation in which the "distribution of earnings is the same for all people with identical innate abilities" (Atkinson, 1980a). Another economist posits that EO requires that "all those who expend the same degree of effort end up with the same outcome, regardless of their circumstances" (Roemer, 2008). A philosophical synthesis suggests that "If there is a broad unifying theme between conceptions of Equality of Opportunity, it is the notion of a justified hierarchy or inequality which classifies some factors as being inappropriate determinants of persons' success" (Elford, 2023).

EO is defined as equal outcomes conditional on effort and ability, once circumstances are held constant. Formally, this implies

$$\frac{\partial \mathbf{C}}{\partial \sigma_i} = 0 \quad \text{for all } i \in \Sigma_C \tag{11}$$

Circumstances refer to morally arbitrary traits—such as family background, race, or gender—for which individuals should not be held accountable. Identifying these traits is a delicate political and philosophical matter, as emphasized in the works of (Roemer, 1998; Dworkin, 2002; Fleurbaey, 2008). This is a normative definition which contrasts sharply with the definition of circumstance, which is positive and a scientific matter.

4.2 Connection among Concepts

The two concepts of EO and IGM, when defined with a focus on outcomes, diverge along two key dimensions. First, EO of outcome holds effort fixed, whereas IGM does not. Definitions 10 and 11 make this distinction explicit: the former involves a total differentiation, while the latter relies on a partial differentiation. The idea of holding effort constant underpins the concept of EO a la Roemer and represents a significant hurdle in its empirical assessment, given the difficulty in observing effort.

Second, EO and IGM consider different sets of traits. IGM centers on hereditary traits, Σ_H, as shown in definition 10, whereas EO focuses on morally arbitrary traits—i.e., circumstances—denoted by Σ_C in definition 11.

The first discrepancy disappears when the focus shifts from outcomes to welfare, since effort is endogenous. To see that, let us start with defining IGM of welfare, as zero correlation in welfare across two generations. Formally, this can be expressed as:

$$Corr(\mathbf{V}(\widetilde{\sigma}), \mathbf{V}(\sigma)) = 0, \tag{12}$$

IGM of welfare is equivalent to equal welfare across all hereditary traits:

$$\frac{\partial \mathbf{V}}{\partial \sigma_i} = 0 \quad \text{for } i \in \Sigma_H. \tag{13}$$

In contrast, EO of welfare requires equal welfare across all circumstances:

$$\frac{\partial \mathbf{V}}{\partial \sigma_i} = 0 \quad \text{for } i \in \Sigma_C. \tag{14}$$

The first distinction between IGM and EO of outcome loses force here, as both definitions involve partial derivatives. In fact, there is no effort to condition on, since effort is endogenous—and we do not condition on an endogenous variable.

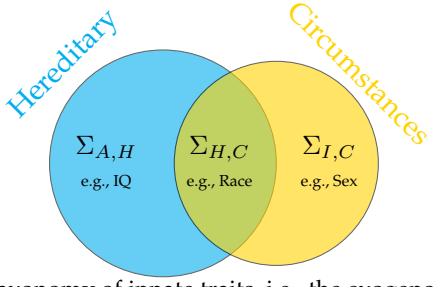
The core distinction between IGM and EO lies solely in the set of traits considered (Swift, 2004; Jencks and Tach, 2006; Torche, 2015). This distinction is illustrated further in the trait space in Figure 3. The outer set of idiosyncratic abilities—traits that are neither hereditary nor circumstances—appears to be a narrow category, potentially including attributes such as footedness. Idiosincratic circumstances includes traits such as sex and De Novo genetic disorders (e.g., Down Syndrome, Congenital Heart Defects, some forms of Autism).

Both IGM and EO are concerned with hereditary circumstances, denoted $\Sigma_{H,C}$—traits like race, place of birth, family connections, or last names. However, EO also accounts for idiosyncratic circumstances ($\Sigma_{I,C}$), such as gender or certain disabilities, which are not necessarily hereditary. This inclusion means that discrimination is inherently part of EO analysis.

Conversely, IGM also considers hereditary abilities ($\Sigma_{A,H}$), such as IQ or height. The latter is an unintended consequence that simplifies the measurement of IGM (Intergenerational Mobility) and explains why perfect mobility is not a desired outcome.

When hereditary traits and circumstances fully overlap—as in a Rawlsian world—EO and IGM of welfare coincide. But if the two sets are distinct, the concepts diverge: higher IGM does not necessarily imply greater EO.

Figure 3: Fundamental IGM & EO Diff: Focus on Different Traits



Notes: This figure presents a taxonomy of innate traits, i.e., the exogenous parameters influencing an individual's decision-making process. The examples provided within each group are illustrative as the definition of circumstances is a normative decision, a political and philosophical matter.

Now, we are prepared to describe the two principal concepts of EO. 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).

EO for outcome requires equal outcome for equal effort, in contrast to EO requiring equal outcome for equal cost.[11] Having established a definition for these concepts, we now turn our attention to exploring their interrelation.

Whether the two concepts are close—and thus whether IGM serves as a reliable proxy for EO—depends on the degree of overlap between hereditary traits and circumstances. Consider a world where the two sets largely coincide, meaning that nearly all circumstances are hereditary and vice versa. In such a world, the roles of hereditary abilities (e.g., genes) and idiosyncratic circumstances (e.g., sex or de novo genetic disorders) are minimal, and the two concepts converge; IGM becomes a good surrogate for EO. In the opposite world, where the key traits are either hereditary or circumstantial, but not both, the two concepts are independent.

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.[12]

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 **??** 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 Figure 3.

[11] To see that, write the individual problem 8 as $V(\sigma) = \max_k U\left(Y\left(K^{-1}(k,\sigma),\sigma\right),k\right)$.

[12] 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

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

5 Conclusion

In an early expression of the career–family tradeoff, in 1492, Cassandra Fedele—a prominent humanist writer—wrote to Alessandra Scala, a Florentine scholar: "And so, my Alessandra, you are uncertain whether to dedicate yourself to the Muses or to a Man?"[13] Both women managed to have both—a career and a family. They married and achieved enough success to be remembered, perhaps thanks to their privileged backgrounds, as argued in this paper.

The concept of Equality of Opportunity is appealing in its aim to level the playing field, yet remains difficult to define and measure. Traditionally, economics has followed two main approaches. The first uses intergenerational mobility as a proxy for equality of opportunity (EO). While it primarily captures equality of outcomes, it offers the advantage of simplicity in measurement. The second defines EO as equal outcomes for equal effort, conditional on circumstances—an approach that hinges on identifying proxies for effort, a complex and often contested task.

This paper proposes a third, novel approach: defining EO as equal outcomes for equal costs. We quantify these costs using data on family investments. We then show that the difference between our welfare-based mobility measure and EO is purely normative.

[13] "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," .

ATKINSON, A. B. (1980a): "Intergenerational income mobility," *JHS Journal*, 3, 61–73.

——— (1980b): "On intergenerational income mobility in Britain," *Journal of Post Keynesian Economics*, 3, 194–218.

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.

BERGER, T., P. ENGZELL, B. ERIKSSON, AND J. MOLINDER (2023): "Social mobility in Sweden before the welfare state," *The Journal of Economic History*, 83, 431–463.

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/.

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*.

——— (2014): "A grand gender convergence: Its last chapter," *American economic review*, 104, 1091–1119.

——— (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," .

JENCKS, C. AND L. TACH (2006): "Would Equal Opportunity Mean More Mobility?" in *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, ed. by S. L. Morgan, D. B. Grusky, and G. S. Fields, Palo Alto: Stanford

University Press, 23–58.

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.

LONG, J. AND J. FERRIE (2013): "Intergenerational occupational mobility in Great Britain and the United States since 1850," *American Economic Review*, 103, 1109–1137.

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.

PLAMENATZ, J. P. (1975): "Equality of Opportunity," in *Aspects of Human Equality*, ed. by L. e. a. Bryson, New York.

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

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

ROEMER, J. E. AND A. TRANNOY (2015): "Equality of opportunity," in *Handbook of income distribution*, Elsevier, vol. 2, 217–300.

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.

SWIFT, A. (2004): "Would Perfect Mobility Be Perfect?" *European Sociological Review*, 20, 1–11.

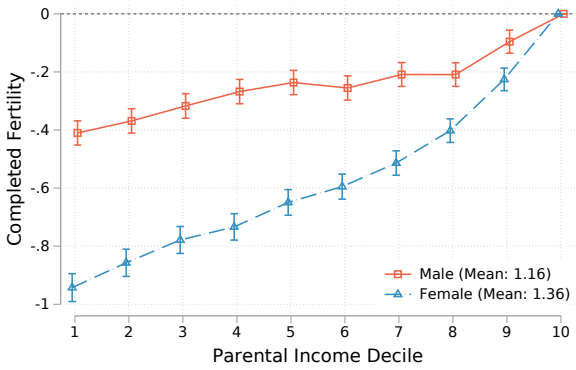
TAUBMAN, P. (1979): "Equality of opportunity and equality of outcome: lessons from studies of twins," *JHS-Journal*.

TAWNEY, R. H. (1952): *Equality*, London, 4th ed.

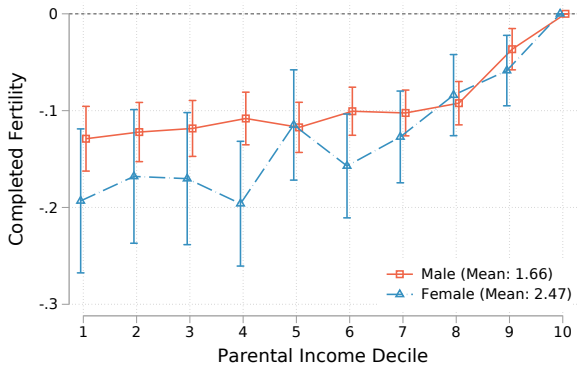
TORCHE, F. (2015): "Intergenerational Mobility and Equality of Opportunity," *European Journal of Sociology*, 56, 343–371.

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

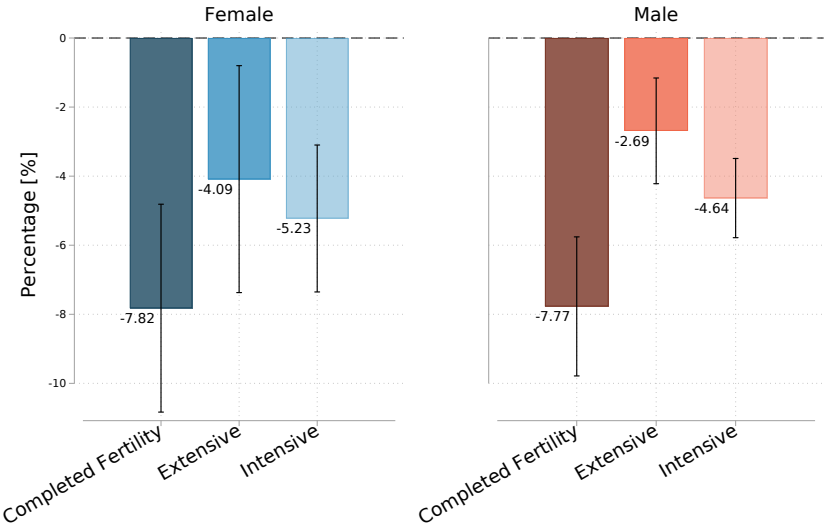
Figure 4: Family Cost of Upward Mobility: Completed Fertility



(a) Linear Specification



(b) Non-parametric Specification



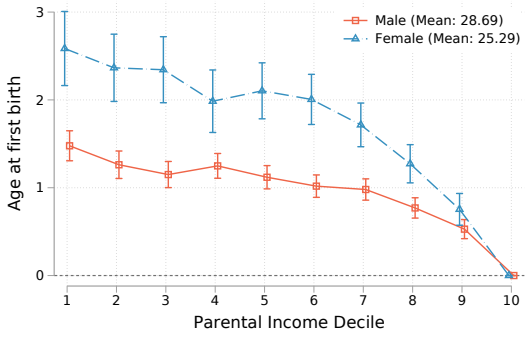
(c) Extensive and Intensive Margin

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, $\{\beta_i\}$, in the regression of family outcome k on a gender-specific non-parametric functions of their own lifetime income c and their parental lifetime income $\tilde{c}$. In Panel 4a, we use the linear specification of the interaction terms described in Equation 6. In Panel 4b, we use the non-parametric specification of interaction terms:

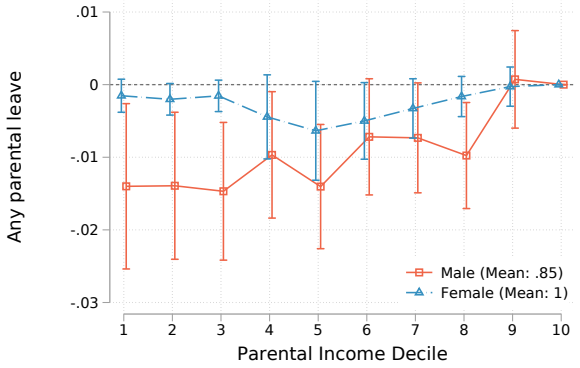
$$k = f(c) + g(\tilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{16}$$

In Panel 4c, we only report β_1 of the above regression, representing the difference in outcomes of disadvantaged achievers from the bottom-decile parents. We normalize the estimate by the baseline mean in Panel 4c. The baseline mean is calculated as the average completed fertility among the bottom-decile-income individuals with bottom-decile-income 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 is the sum of labor incomes between ages 25 and 45. Parental lifetime income $\tilde{c}$ is the sum of labor, business, and capital income between the individual's ages 0 and 20. Completed fertility includes all biological offspring born before age 45.

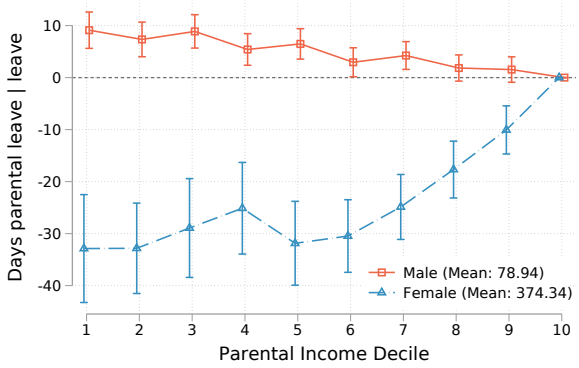
Figure 5: It's All About Time: Timing of Birth and Parental Leave



(a) Age at First Childbirth



(b) Parental Leave per Child: Extensive Margin



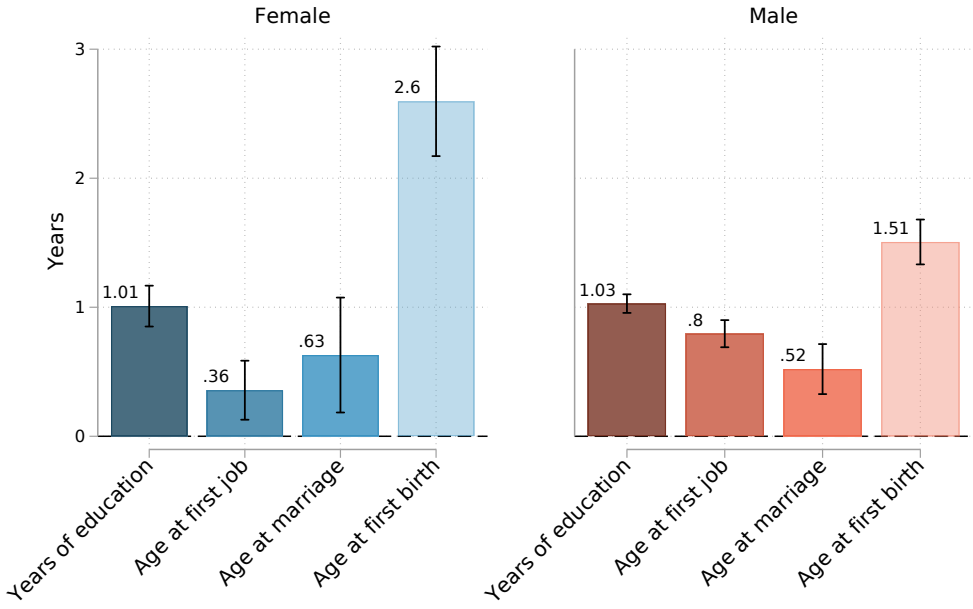
(c) Parental Leave per Child: Intensive Margin

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, $\{\beta_i\}$, in the regression of family outcome k on a gender-specific non-parametric functions of their own lifetime income c and their parental lifetime income $\tilde{c}$:

$$k = f(c) + g(\tilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{18}$$

The baseline mean is calculated as the average of the outcome variable among the bottom-decile-income individuals with bottom-decile-income 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 5b and Panel 5c further restrict the sample to birth cohorts of 1969-1975 for whom we observe parental leave days by age 45. Own lifetime income c is the sum of labor incomes between ages 25 and 45. Parental lifetime income $\tilde{c}$ is the sum of labor, business, and capital income between the individual's ages 0 and 20. Completed fertility includes all biological offspring born before age 45.

Figure 6: Time Drawbacks: Length of Studies, Start in the Labor Market, Start of a Family



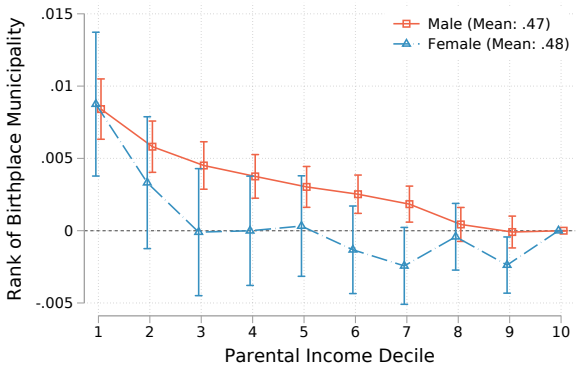
(a) Age at starting Family and Professional Life

Note: This figure illustrates the difference in family and career 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 coefficient of the interaction term, $\{\beta_1\}$, in the regression of family outcome k on a gender-specific non-parametric functions of their own lifetime income c and their parental lifetime income $\tilde{c}$:

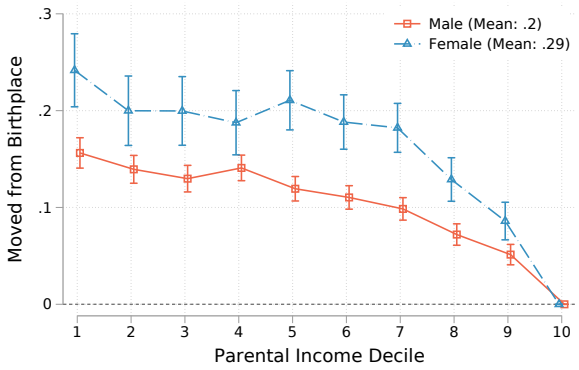
$$k = f(c) + g(\tilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{20}$$

Each bar reports β_1 running the above regression on the dependent variable of interest. 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. *Age at marriage* is conditional on being married before the age of 45. *Age at first child* is conditional on giving birth before age 45. More details are provided in Section 2.

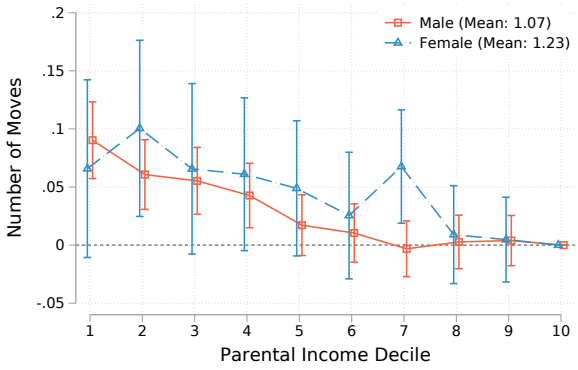
Figure 7: Location Drawbacks: Geographical Mobility



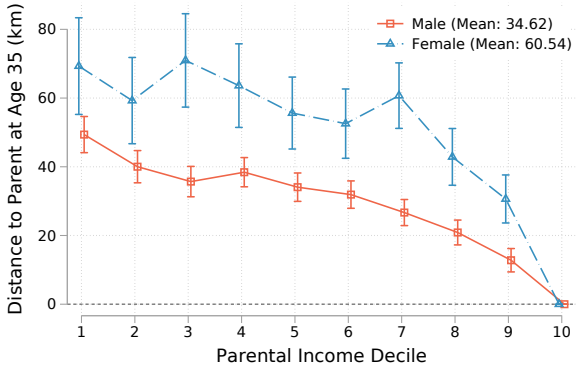
(a) Birthplace Income Rank



(b) Moved From Birthplace



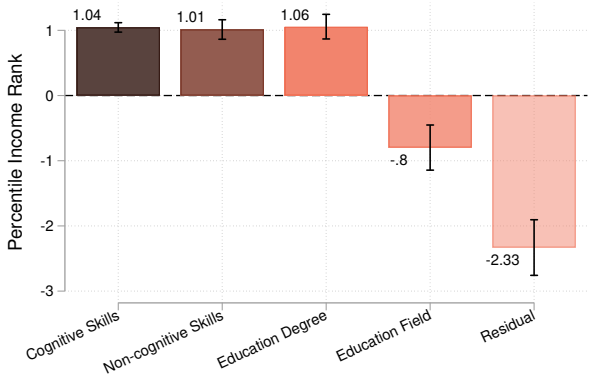
(c) Number of Moves



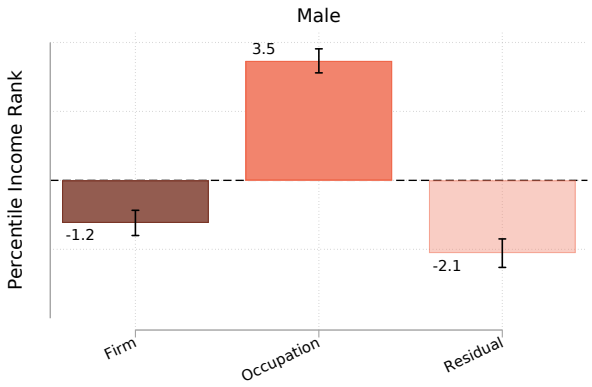
(d) Distance to Parent

Note: This figure illustrates the relative different careers of individuals reaching the top-decile of lifetime income as a function of their parental income deciles ($\{\beta_l\}$ of Equation 6). 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 a parent's residence address. We use the minimum distance when parents have different residences. 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 bottom-decile-earning individuals with bottom-decile-earning parents. For more details, see Section 2.

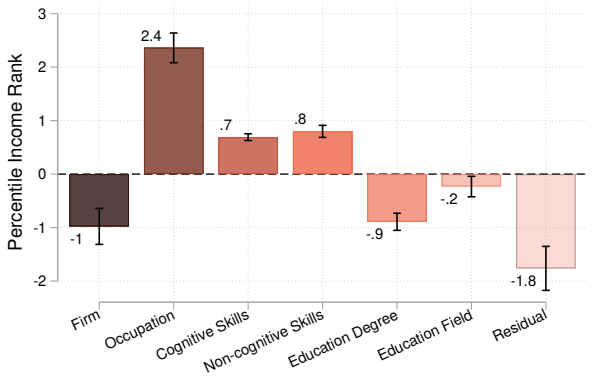
Figure 8: Mechanism



(a) Human Capital Inequality



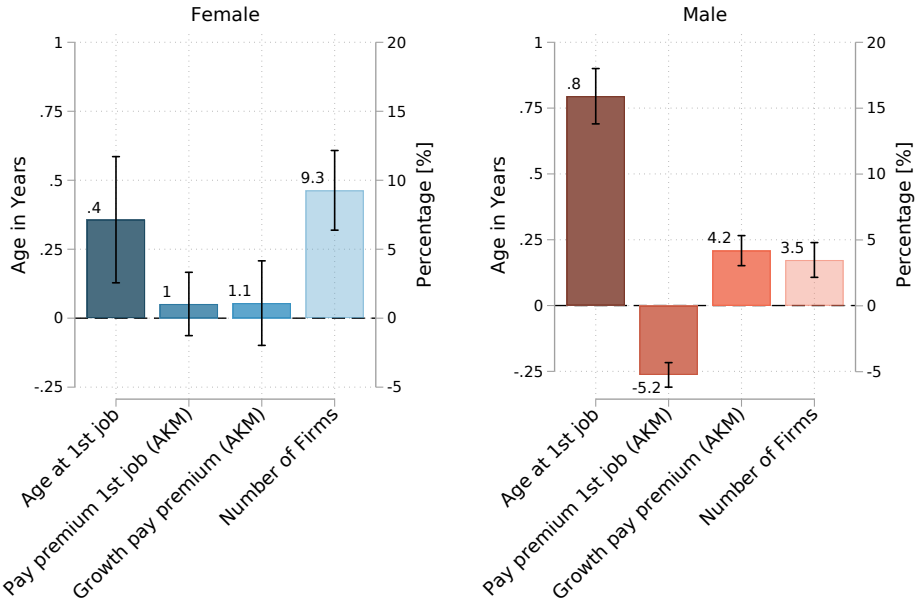
(b) Labor Market Inequality



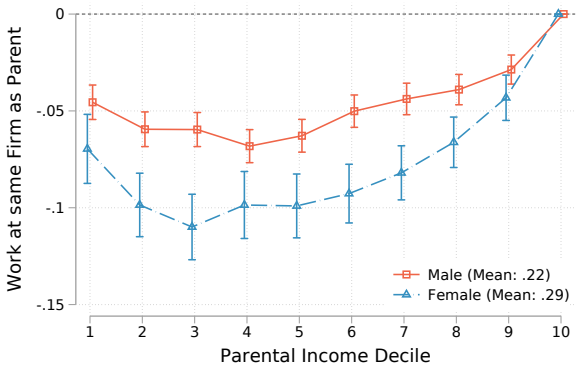
(c) Labor Market Inequality conditional on Human Capital

Note: This figure illustrates the relative difference in human capital and labor market outcomes between individuals reaching the top decile of lifetime income from different parental income deciles. It plots the coefficient $\{\beta_1\}$ in Regression 7 where outcomes are four separate measures of human capital and two separate measures of the labor market. The measures in each figure are harmonized using their predicted lifetime income ranks. Specifically, we use the predicted values of each measure from a Mincerian regression of lifetime income rank on human capital measures (Panel 8a), on labor market measures (Panel 8b), and the combination of the two (Panel 8c). 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.3. 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, 336 categories of field of study, and 123 occupation categories (see Appendix Tables A.2, A.3, and A.4).

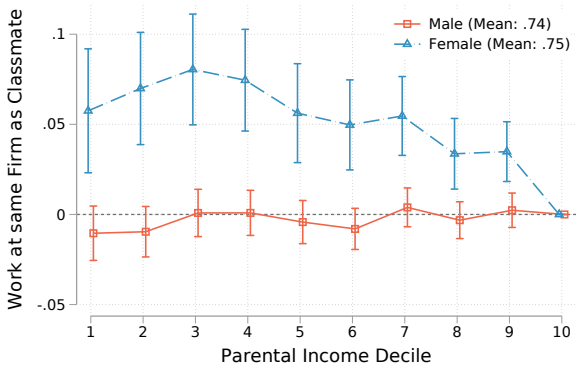
Figure 9: Achievers from Underprivileged Backgrounds are Late Bloomers



(a) Differences in Career Trajectories



(b) Same Firm as Parent



(c) Same Firm as Classmate

Note: This figure illustrates the relative different career outcomes of individuals reaching the top-decile of lifetime income as a function of their parental income deciles ($\{\beta_l\}$ of Equation 7). Panel 9a reports $\{\beta_1\}$ for different type of career measures. *Firm Pay Premium* is measured as the AKM firm fixed effect in the fashion of Abowd et al. (1999); Card et al. (2013). *Number of firms* 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. 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 9b and Panel 9c illustrate the differences in access to networks. We measure the outcome as having worked in the same firm between ages 30 and 45 as any firm their parents or their university peers have ever worked for. We define university peers as the group of individuals who graduate in the same year, the same university major, and the same municipality. The baseline mean reports the average of dependent variables among the bottom-decile-earning individuals with bottom-decile-earning parents. For more details, see Section 2.

Table 1: Sample Characteristics

	Mean	SD	Min	Max
Female	0.48	(0.50)	0.00	1.00
Year of Birth	1965.14	(6.01)	1955.00	1975.00
Number of siblings	1.90	(1.25)	0.00	7.00
Education >= college	0.34	(0.47)	0.00	1.00
Completed Fertility	1.90	(1.12)	0.00	5.00
Any Children	0.85	(0.36)	0.00	1.00
Days of parental leave per child	214.04	(182.58)	0.00	700.00
Age at first child ⌈ child before age 45	28.68	(5.48)	17.00	44.00
Years of Education	12.51	(2.26)	7.00	20.00
Age at first Employment	26.02	(5.38)	20.00	65.00
Age marriage ⌈ married before age 45	31.03	(5.68)	18.00	45.00
Rank of Birthplace Municipality	0.50	(0.06)	0.27	0.63
Moved from Birthplace	0.31	(0.46)	0.00	1.00
Number of Moves between age 35-45	0.89	(1.17)	0.00	10.00
Distance to Parent at Age 35 (km)	67.23	(140.90)	0.00	1442.91
Number of Firms between Age 30-45	4.16	(1.90)	0.00	17.00
Work at same Firm as Parent between Age 30-45	0.22	(0.42)	0.00	1.00
Work at same Firm as Classmate between Age 30-45	0.72	(0.45)	0.00	1.00
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 2: Overrepresented Occupations Among Top-Decile Individuals Relative to Population

Top Decile Parents	Overrep.	**Bottom Decile Parents**	Overrep.
Chief Executives	10.04	Pilots	6.08
Health/Medical Specialists	8.80	Senior Officials/Politicians	5.91
Lawyers	7.02	Specialized Managers	5.69
Specialized Managers	6.93	Chief Executives	5.34
Senior Officials/Politicians	6.77	Military	4.46

Note: This table shows the top five overrepresented occupations for advantaged (disadvantaged) achievers in the left (right) column. Overrepresentation is measured as the share of a given occupation within the group divided by its share in the overall sample.

Online Appendix of:

Invisible Cost of Upward Mobility

Patrizia Massner & Arash Nekoei

April 17, 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

$$\tau\left(Z,\widetilde{Z}\right) = 4\mathbb{E}\left(C_{Z,\widetilde{Z}}\left(Z,\widetilde{Z}\right)\right) - 1 \tag{21}$$

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

$$\begin{aligned}\tau\left(Z,\widetilde{Z}\right) &= \mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) > 0\right) - P\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) < 0\right) \\ &= 2\mathbb{P}\left((Z_2 - Z_1)\left(\widetilde{Z}_2 - \widetilde{Z}_1\right) > 0\right) - 1\end{aligned} \tag{22}$$

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

$$\tau\left(Z,\widetilde{Z}\right) = 4\mathbb{P}\left(Z_1 < Z_2, \widetilde{Z}_1 < \widetilde{Z}_2\right) - 1 \tag{23}$$

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 23 and 21 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

$$\tau\left(Z, \widetilde{Z}\right) = 1 - 4\chi\left(Z, \widetilde{Z}\right) \tag{24}$$

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 24, that is

$$\tau_g = \tau\left(g(\Gamma, \Theta), g\left(\widetilde{\Gamma}, \widetilde{\Theta}\right)\right) = 1 - 4\,\chi_g \tag{25}$$

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

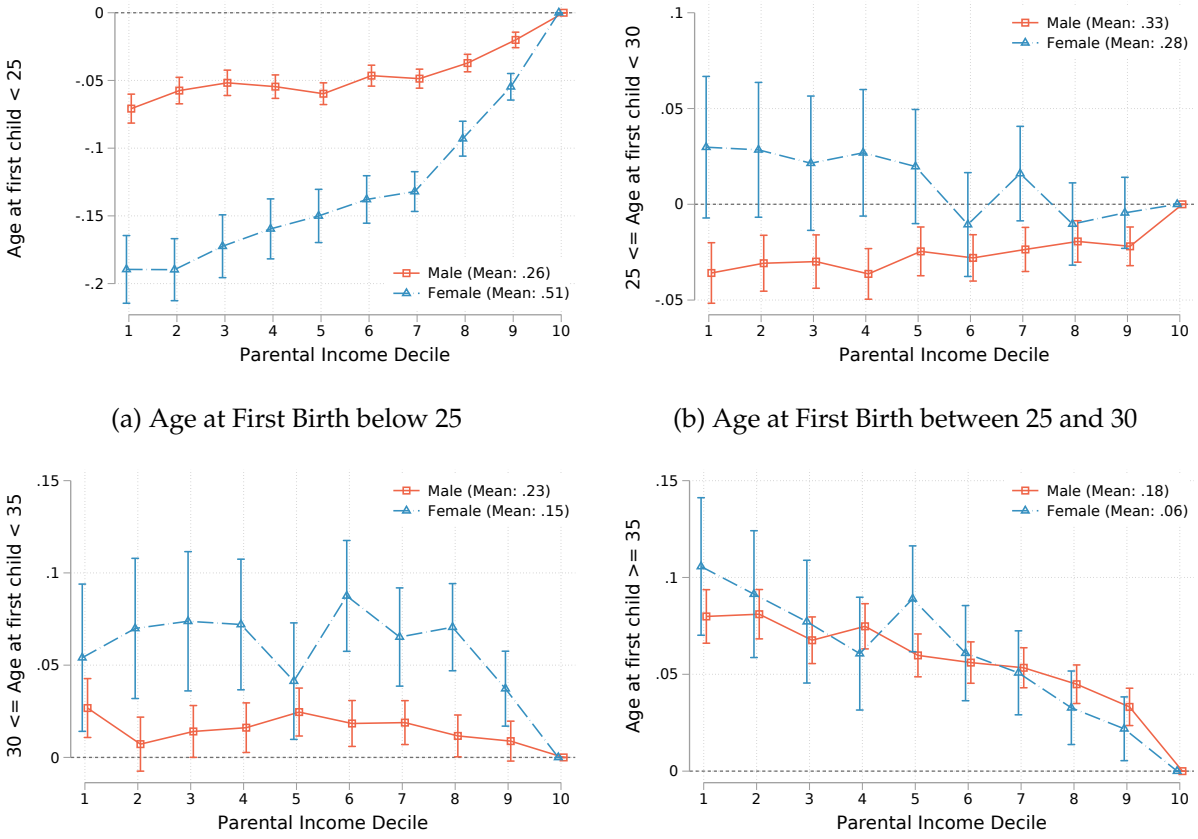
The Proposition implies that $\frac{d}{d\gamma}\frac{dK/d\theta}{dC/d\theta} < 0 \Leftrightarrow \alpha_k > \alpha_c$. We show below $\alpha_k > \alpha_c$ is equivalent to higher MRS for income than welfare. To show that $\alpha_k > \alpha_c$ is equivalent to higher MRS for income than welfare, start 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 the lemma that showed that higher MRS implies strong strict mobility order. $Q.E.D.$

B Appendix Figures

Figure A.1: 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

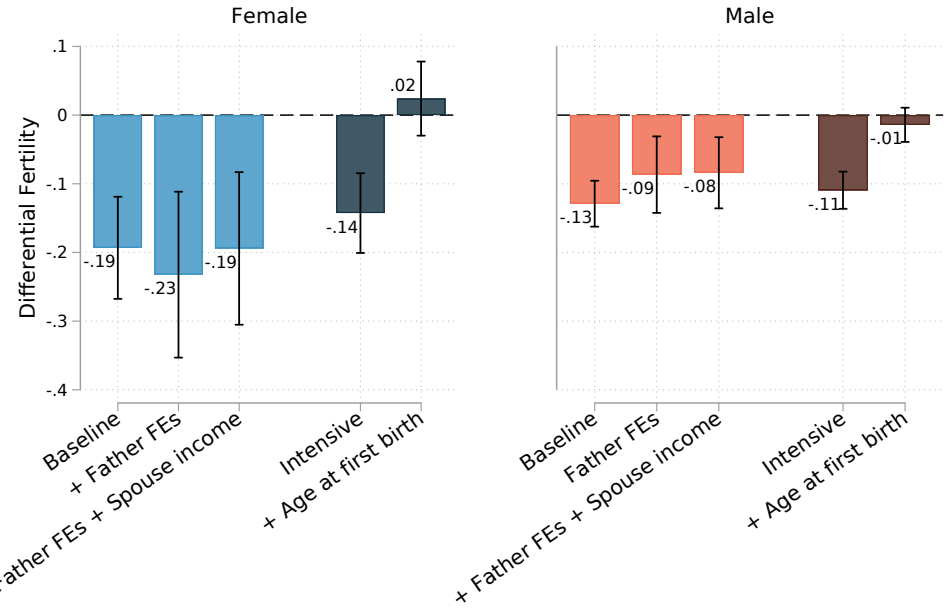
(d) Age at First Birth above 35

Note: This figure illustrates the difference in the age at the first birth 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, $\{\beta_i\}$, in the regression of family outcome k on a gender-specific non-parametric functions of their own lifetime income c and their parental lifetime income $\tilde{c}$:

$$k = f(c) + g(\tilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\tilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{27}$$

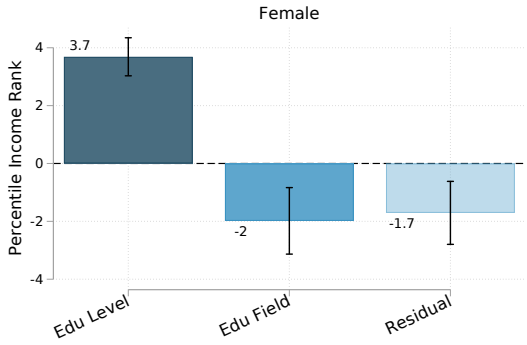
The baseline mean is calculated as the average of the outcome variable among the bottom-decile-income individuals with bottom-decile-income parents. The estimation sample is based on 1,356,400 individuals born in Sweden between 1955 and 1975 from parents who were also born in Sweden. All family outcomes in this figure are conditional on having at least one child. Own lifetime income c is the sum of labor incomes between ages 25 and 45. Parental lifetime income $\tilde{c}$ is the sum of labor, business, and capital income between the individual's ages 0 and 20.

Figure A.2: Adding controls

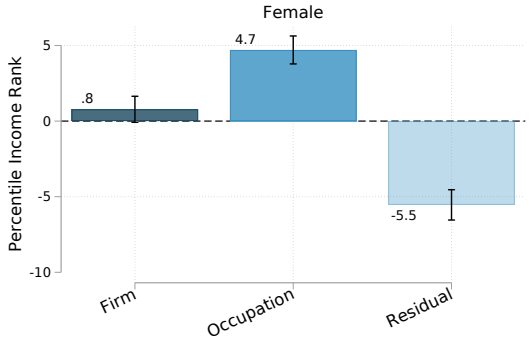


Note: This figure shows the relative difference in the number of children between individuals at the top vs. the bottom of lifetime income for individuals with bottom vs. top parental income, where top decile parents are the baseline group. It plots the coefficients $\{\beta_1\}$ in Regression 7 where we additionally add control variables to the main regression as indicated. 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. Adding father fixed effects requires to restrict the sample to individuals with at least one sibling. Adding the age at first birth as control variable requires to restrict the sample to individuals who have at least one child.

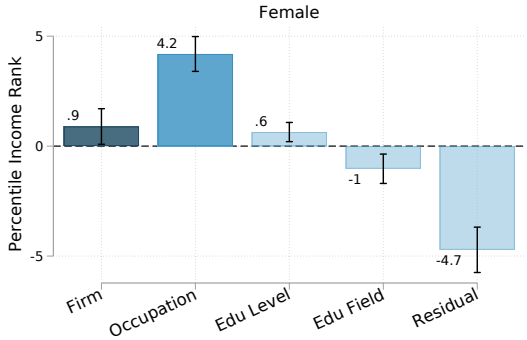
Figure A.3: Mechanism: Females



(a) Human Capital Inequality



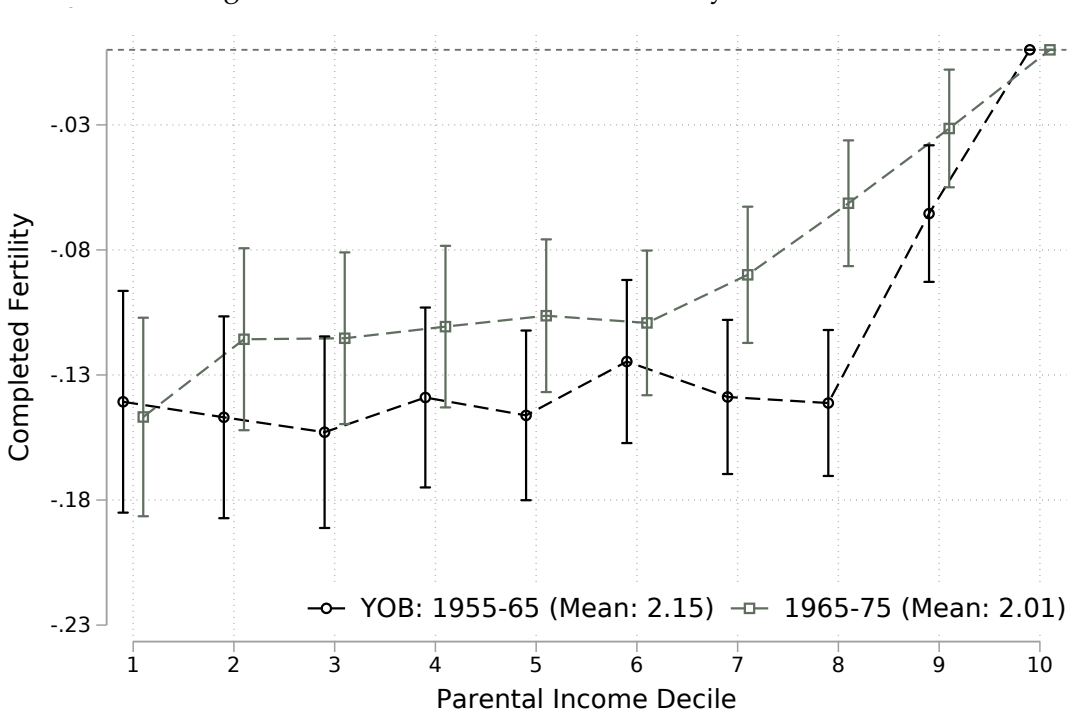
(b) Labor Market Inequality



(c) Labor Market Inequality conditional on Human Capital

Note: This figure reports the corresponding results for females, using a similar design as Figure **??**. Cognitive and non-cognitive skill measures are omitted, as these are only available for men in the military draft data. This figure illustrates the relative difference in human capital and labor market outcomes between individuals reaching the top decile of lifetime income from different parental income deciles. It plots the coefficient $\{\beta_1\}$ in Regression 7 where outcomes are two separate measures of human capital and two separate measures of the labor market. The measures in each figure are harmonized using their predicted lifetime income ranks. Specifically, we use the predicted values of each measure from a Mincerian regression of lifetime income rank on human capital measures (Panel A.3a), on labor market measures (Panel A.3b), and the combination of the two (Panel A.3c). We exploit 47 distinct categories of highest completed education level, 336 categories of field of study, and 123 occupation categories (see Appendix Tables A.2, A.3, and A.4).

Figure A.4: Evolution of Cost of Mobility over Time

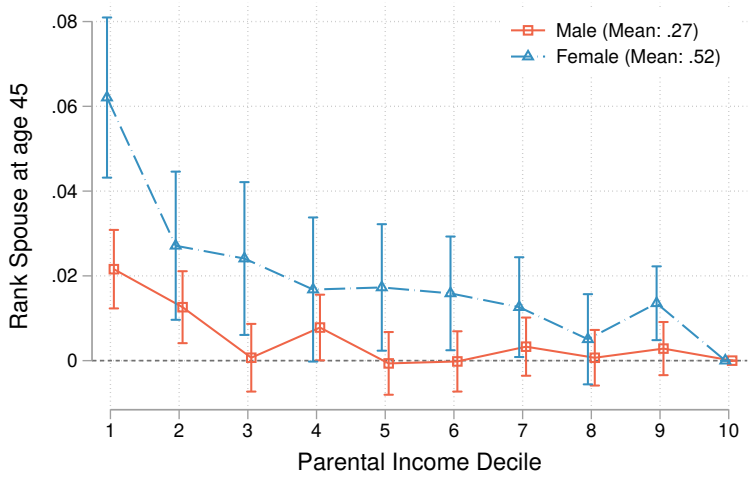


Note: This figure shows the coefficient estimates of running the following regression separately by birth cohort groups:

$$k = f(c) + g(\widetilde{c}) + \sum_{i=1,\ldots,9} \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\widetilde{c}) = i)}_{\text{Parental Income Deciles}} + \varepsilon \tag{29}$$

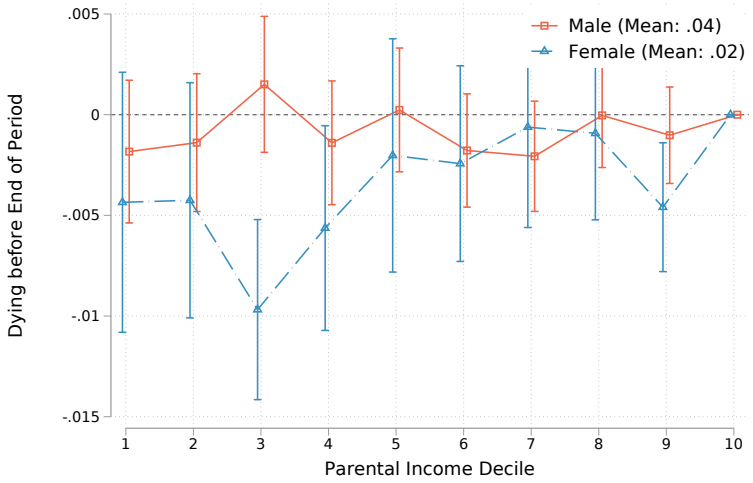
While we interact the control variables $f()$ and $g()$ by gender, we run the interaction term of interest pooled by gender. We categorize individuals into two groups based on their year of birth (YOB): Those born between 1955 and 1965 (black line) and those born between 1965 and 1975 (green line). The 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 c is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income $\widetilde{c}$ is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. The number of children includes all biological offspring born before age 45. For more details see Section 2.

Figure A.5: 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.

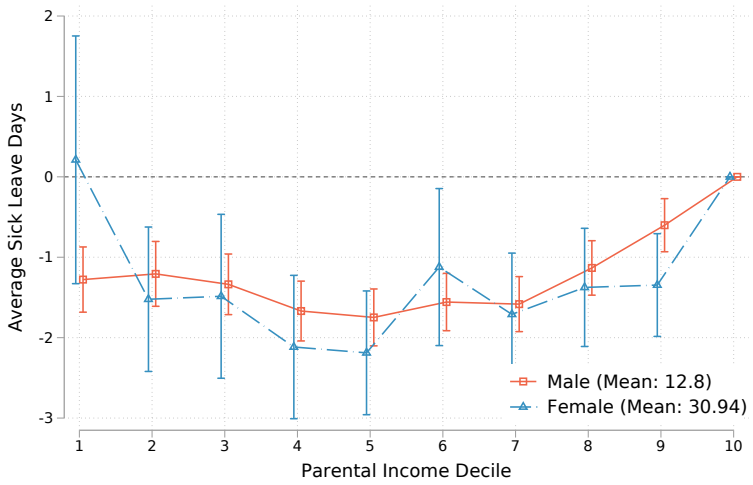
Figure A.6: Mortality



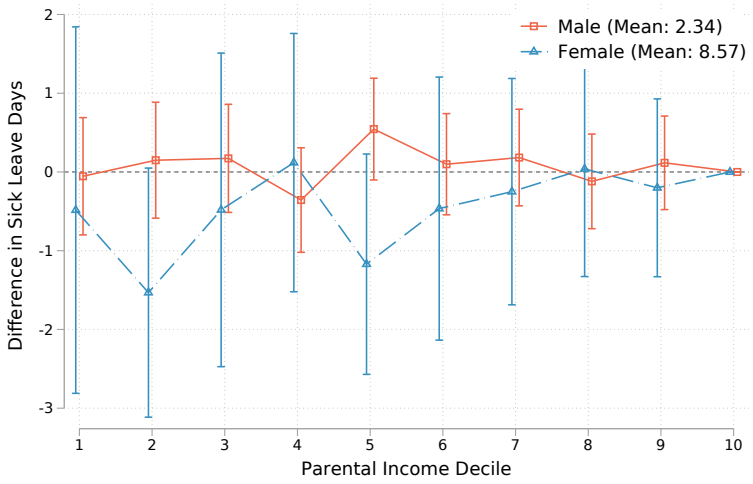
(a) Interaction, β_l

Note: This figure 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 individuals' outcome k on their own lifetime income c and their parental (father + mother) lifetime income $\widetilde{c}$ where we interact the regression by gender. We additionally add controls of year and month of birth to the regression. The outcome variable, 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 point in time we observe the 1970-cohort before the data ends in year 2020. Own lifetime income is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income is measured as the sum of the father's and mother's total income over the individual's ages 0 to 20. 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 bottom-decile income bracket with parents in the bottom-decile income bracket. For more details see Section 2.

Figure A.7: Days of Sick Leave



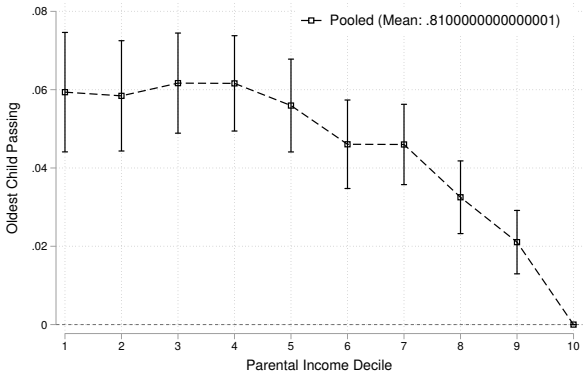
(a) Level



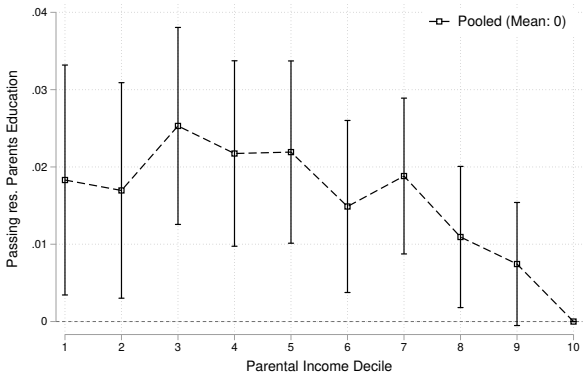
(b) Change

Note: This figure 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 on individuals' own lifetime income c and their parental (father + mother) lifetime income $\widetilde{c}$, where we interact the regression by gender. We additionally control for year of birth and month in the regression. Due to data limitations, we compute the average number of sick leave days per year between age 35 and 45. The outcome variable, *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. Own lifetime income is measured as the sum of all labor income between ages 25 and 45. Parental lifetime income 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 bottom-decile income bracket with parents in the bottom-decile income bracket. For more details see Section 2.

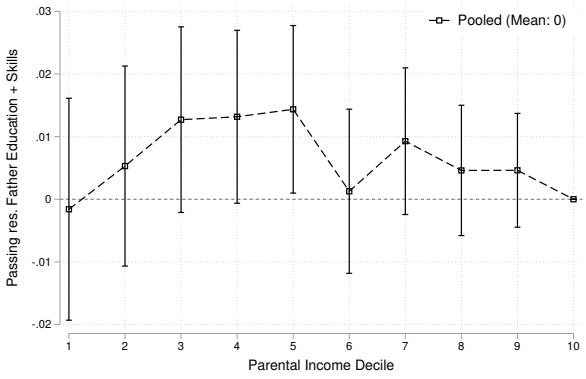
Figure A.8: Children's Educational Achievements: Passing All Subjects



(a) Children's Education



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



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

Note: This figure 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. Specifically, it plots the coefficients of the interaction terms, $\{\beta_i\}$, in the regression of outcome k on individuals' own lifetime income c and their parental (father + mother) lifetime income $\tilde{c}$ where we interact the control variables by gender. The outcome variable, 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. In Panel A.8b we residualize the dependent variable by first regressing the child's school performance on both the father's and mother's education degree and field of study. In Panel A.8c we residualize the dependent variable by first regressing the child's school performance additionally on the father's cognitive and non-cognitive skills (besides education degree and the field of study). We use parents in the top income decile as the reference group. 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 in the national test exams. We report the mean for individuals in the bottom-decile income bracket with parents in the bottom-decile income bracket.

C Appendix Tables

Table A.1: Health Outcomes

	(1) Mortality	(2) Days of 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 (Bottom 10% x Parents Bottom 10%)	.03	22.14
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 7 where we bundle the interaction term and pool by gender:

$$k = f(c) + g(\widetilde{c}) + \beta_i \underbrace{\mathbb{1}(D(c) = 10)}_{\text{Own Income Top Decile}} \times \underbrace{\mathbb{1}(D(\widetilde{c}) \in \{1, ..., 9\})}_{\substack{\text{Parental Income} \\ \text{Bottom 90}}} + \varepsilon \tag{31}$$

where k 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.2: 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). Field of study classifications were originally provided in Swedish. For the purposes of analysis and presentation, all translations into English were carried out by the authors. Any misinterpretations or inaccuracies are the authors' responsibility.

Table A.3: SUN2000 - 3 digit education degree 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 degree codes (47 categories). Education degree classifications were originally provided in Swedish. For the purposes of analysis and presentation, all translations into English were carried out by the authors. Any misinterpretations or inaccuracies are the authors' responsibility.

Table A.4: 3-digit occupation codes

Code	Description (English Translation)
111	Senior officials and politicians
112	Senior officials in interest organizations
121	Chief executives and managing directors
122	Production and operations managers
123	Managers of specialized functions
131	Managers of small enterprises and units
211	Physicists, chemists, and related professionals
212	Mathematicians and statisticians
213	IT specialists
214	Civil engineers, architects, and related professionals
221	Specialists in biology, agriculture, forestry, etc.
222	Health professionals (except nursing)
223	Midwives; nurses with specialized competence
231	University and higher education teachers
232	Upper secondary school teachers
233	Primary school teachers
234	Special education teachers
235	Other teaching professionals with theoretical specialist competence
241	Economists, marketers, and HR officers
242	Legal professionals
243	Archivists, librarians, and related professionals
244	Social scientists and linguists (except social workers)
245	Journalists, artists, actors, and related professionals
246	Clergy
247	Public administration professionals
248	Administrators in interest organizations
249	Psychologists, social workers, and related professionals
311	Engineers and technicians
312	Computer technicians and operators
313	Photographers; audio/video technicians, hospital techs, etc.
314	Pilots, ship officers, and similar
315	Safety and quality inspectors
321	Agricultural and forestry technicians
322	Physiotherapists, dental hygienists, etc.
323	Nurses
324	Biomedical analysts
331	Preschool teachers and leisure-time pedagogues
332	Other teachers and instructors
341	Salespersons, buyers, brokers, etc.
342	Agents, intermediaries, etc.
343	Accountants, administrative assistants, etc.
344	Customs, tax, and social insurance officials
345	Police officers
346	Treatment assistants, recreation leaders, etc.
347	Illustrators, entertainers, professional athletes, etc.
348	Pastors and other religious workers
411	Office secretaries and data entry clerks

412	Bookkeeping and accounting assistants
413	Warehouse and transport clerks
414	Library assistants and similar
415	Mail carriers and sorters
419	Other office staff
421	Cashiers, tellers, and related clerks
422	Customer service clerks
511	Travel attendants and related workers
512	Institutional kitchen and restaurant staff
513	Care and welfare workers
514	Hairdressers and other personal service workers
515	Security personnel
521	Fashion and other models
522	Retail salespersons and demonstrators
611	Crop growers in agriculture and horticulture
612	Animal breeders and handlers
613	Mixed crop and livestock producers
614	Forestry workers
615	Fishers and hunters
711	Miners, quarry workers, stonecutters
712	Building and construction workers
713	Building trades workers
714	Painters, varnishers, chimney sweeps, etc.
721	Metal casters, welders, sheet metal workers, etc.
722	Blacksmiths, toolmakers, etc.
723	Machine and engine mechanics
724	Electrical, telecom, and electronics repairers
731	Precision mechanics and similar trades
732	Potters, glassworkers, decorative painters, etc.
733	Craftworkers in wood, textiles, leather, etc.
734	Graphic arts and related trades workers
741	Food processing workers
742	Furniture and model carpenters
743	Tailors, cutters, upholsterers, etc.
744	Tanners, leather workers, shoemakers
811	Ore processing operators, well drillers, etc.
812	Operators in steel and metal production
813	Operators in glass and ceramic production
814	Operators in wood and paper industries
815	Operators in the chemical industry
816	Power plant operators and related workers
817	Industrial robot operators
821	Machine operators in metal and mineral processing
822	Machine operators in chemical industries
823	Rubber and plastic product machine operators
824	Machine operators in the wood industry
825	Machine operators in printing and paper industries
826	Textile, fur, and leather machine operators
827	Food product machine operators
828	Assemblers

829	Other machine operators and assemblers
831	Locomotive drivers and related workers
832	Motor vehicle drivers
833	Agricultural and mobile machinery operators
834	Deck crew and related ship workers
911	Street and market vendors
912	Cleaners and helpers
913	Kitchen and restaurant assistants
914	Newspaper distributors, caretakers, etc.
915	Sanitation and recycling workers
919	Other service workers
921	Helpers in agriculture, gardening, forestry, and fishing
931	Construction laborers
932	Manual packers and other factory workers
933	Freight handlers and couriers
011	Armed forces
	Missing categories

Note: The table reports the descriptions of the 3 digit occupations. Occupational classifications were originally provided in Swedish. For the purposes of analysis and presentation, all translations into English were carried out by the authors. Any misinterpretations or inaccuracies are the authors' responsibility.