IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16273

Careers and Intergenerational Income Mobility

Catherine Haeck
Jean-William Laliberté

JUNE 2023

DISCUSSION PAPER SERIES

IZA DP No. 16273

Careers and Intergenerational Income Mobility

Catherine Haeck
Université du Quebec à Montréal

Jean-William Laliberté
University of Calgary and IZA

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Careers and Intergenerational Income Mobility*

This paper uses Census microdata linked with tax records to quantify the contribution of career choices – occupations and elds of study – to intergenerational income mobility. We document substantial segregation into occupations by parental income. Yet, the occupations children pursue explain only a third of the intergenerational persistence of income. We further describe patterns of intergenerational occupational following and show they vary substantially across occupations, with low-paying occupations showing more persistence across generations on average. However, clustering into occupations based on parental income is mostly independent of parental occupations. Our results demonstrate that occupational persistence only weakly contributes to income immobility.

JEL Classification: J62, J24

Keywords: intergenerational mobility, occupational choice, income inequality

Corresponding author:
Jean-William Laliberté
University of Calgary
2500 University Drive N.W.
T2N 1N4 Calgary, Alberta
Canada

E-mail: jeanwilliam.lalibert@ucalgary.ca

* We gratefully acknowledge funding from the Social Sciences and Humanities Research Council (SSHRC).

Do career choices drive intergenerational income mobility? Parents often "pass" their occupations on to their children (Laband and Lentz, 1983; Long and Ferrie, 2013; Doepke and Zilibotti, 2017; Lo Bello and Morchio, 2022), producing occupational persistence across generations. Prior work has evaluated the degree of persistence for specific or subsets of occupations, but the economic significance of this phenomenon for overall income mobility remains unclear.

This paper analyzes the importance of career choices – occupations and fields of study – for intergenerational income mobility in the Canadian population. We use administrative tax data linked with confidential Census microdata to fully characterize the intergenerational transmission of occupations and quantify its role for income mobility.

Our analysis proceeds in two steps. First, we measure the extent to which the occupations that children pursue contribute to intergenerational income mobility. The explanatory power of occupations depends on the degree of segregation by parental income across children's occupations, as well as on the distribution of economic returns to specific occupations conditional on parental income. The first part of this paper presents these statistics and performs a formal decomposition analysis of child income gaps across parental income ranks.

Second, we investigate the role of occupational persistence across generations as an explanation for occupational segregation by parental income. Children of high-income parents may end up in high-paying occupations because their parents worked in those same or similar occupations. For instance, prior work has documented the intergenerational transmission of specific occupations such as physicians (Lentz and Laband, 1989), lawyers (Laband and Lentz, 1992; Raitano and Vona, 2021), pharmacists (Mocetti, 2016), as well as liberal professions (Aina and Nicoletti, 2018) and professional services (Mocetti et al., 2020) more generally. We extend the analysis to encompass all 500 occupations listed in the Canadian census. The granularity of occupational codes combined with large sample sizes allow us to compute the entire transition matrix linking children's occupations to their parents' occupations, thereby accounting non-parametrically for both direct (e.g. children of physicians becoming physicians too) and off-diagonal occupational linkages (e.g. children of dentists becoming physicians). The latter may be important if parental occupations affect the likelihood of children pursuing a *similar* occupation that requires a related skill set, but not necessarily the same occupation as their parents.

Methodologically, we adopt a rank-rank specification to characterize intergenerational income mobility (Dahl and DeLeire, 2008; Chetty et al., 2014; Deutscher and Mazumder, 2020), which lends itself naturally to between-group decomposition methods since it partitions the sample into income-based groups. Our analyses are based on a generalized version of regression-compatible Kitagawa-Oaxaca-Blinder-type decomposition approaches (Fortin, 2008). This permits the explanatory power of career variables (occupations or fields of study) to vary flexibly across the parental income distribution. In addition, this discretization of parental income into groups allows us to assess the importance of occupational transmission transparently by using simple re-weighting procedures. That is, we obtain counterfactual distributions of children's occupations under a scenario in which we fix transition paths between parents' and children's occupations across parental income groups (i.e. imposing occupational transition probabilities conditional on parental occupations to be the same for everyone). We thereby isolate differences across parental income groups that can be traced back to differences in the parental occupations they inherit, fixing across groups how these occupations are passed on.

There are three main sets of findings. First, we document considerable segregation by parental income across children's occupations. For instance, 53% of physicians come from the top quintile of the parental income distribution, whereas only 6% come from the bottom quintile. In contrast, 10% of light-duty cleaners come from the top quintile of the parental income distribution, and 28% from the bottom quintile. These patterns of segregation are asymmetric, with high-paying occupations being more segregated along parental income lines than low-paying ones.

Second, the occupations children pursue explain roughly one third of the total intergenerational transmission of income. That is, parental income still has important predictive power even among children working in the same occupations. Most of the variance in income occurs within rather than across occupations, and parental income correlates strongly with residual within-occupation child income. We note that the explanatory power of occupations for income persistence is greater for women than for men, in part because men are dramatically over-represented in high upward mobility occupations – high returns occupations that disproportionately draw children from low-income families – in the mining and construction industries.

Third, we find that having a parent in a given occupation increases the likelihood of pursuing that occupation for almost all occupations, but there is considerable heterogeneity. For example, while children of plumbers are 30 times more likely to be plumbers than the rest of the population, children of financial analysts are about 4 times more likely to be financial analysts, despite the total number of children in these two occupations being roughly the same. Yet, only a small fraction – about 10% – of the total intergenerational income rank-rank relationship is accounted for by such patterns of occupational transmission. Intuitively, while the conditional probabilities of pursuing a given occupation if one's parent works in that occupation are large, the base rates are very small. For example, children of specialist physicians are 23 times more likely to become specialist physician than others, but since few children have specialist physician parents, the vast majority of specialist physicians do *not* have specialist physician parents. As a result, the explanatory power of occupational following for total income mobility remains limited.

Overall, we find that children of high-income families do cluster into high-paying occupations, but that intergenerational occupational following plays a limited role explaining this phenomenon. Segregation into occupations is primarily based on socioeconomic status, not persistence of occupations within families. In addition, there is a strong relationship between child and parent income even among children working in the exact same occupation. That is, children from low-income families who successfully enter high-paying occupations earn less on average than children from high-income families who work in those same occupations. In complementary analyses we rule out the intergenerational transmission of occupation-specific human capital within the family as an explanation for this within-occupation income gradient.

This paper bridges two streams of literature on intergenerational income mobility and on occupational persistence. On one hand, recent innovations in income mobility research made possible by the availability of administrative tax data (e.g., Corak and Heisz (1999), Chetty et al. (2014, 2016) and Connolly et al. (2022)) have shown that mobility is correlated with family structure, parental education, as well as inequality and segregation by income and race. The mechanisms through which income is transmitted across generations, however, remain to be better understood. On the other hand, sociologists have long been interested in the transmission of social class and occupations from parents to children (e.g. Erikson et al. (1979); Blau and Duncan (1967)). Our analyses show that, in an accounting sense, occupational following accounts for very little of total intergenerational income transmission. Children of high-income parents do pursue higher-paying occupations, but occupational following only explains a small fraction of this phenomenon. Our work also indirectly speaks to measurement issues in studies of intergenerational mobility. For instance, it is common to use occupational status for the study of intergenerational mobility when historical data do not include information on income (Jácome et al., 2021; Mazumder and Acosta, 2015; Long and Ferrie, 2013; Modalsli, 2017). However, income mobility can differ markedly from occupational or class mobility, as emphasized by Björklund et al. (2000), Blanden (2013) and Torche (2015). In this paper, we consider the role of career choices as a mechanism through which income is transmitted across generations. Consistent with prior findings that occupational and income mobility are empirically distinct, we find that for recent birth cohorts in Canada, two-thirds of the income rank-rank relationship is unrelated to occupational status. Inputting income on the basis of occupations will therefore lead analysts to severely mischaracterize the persistence of income across generations.

1 Data

We use data from Statistics Canada's Intergenerational Income Database (IID) combined with de-identified Canadian Census microdata. The IID includes administrative tax data covering fiscal years 1978 to 2016 inclusively for all Canadians born between 1963 and 1985 (except for those born in 1971, 1976 and 1981) as well as for their parents.[1] These longitudinal fiscal files are linked with six waves of the long-form Census (1991, 1996, 2001, 2006, 2011, 2016), each of which are filled by random samples of 1 out of 5 Canadian households.

Sample Selection Our main sample consist of all children in the IID that (1) are matched to at least one Census wave, and (2) have at least one parent matched to at least one Census wave. It includes 31% of all children included in the IID, representing about 2.9 million children.[2] One might be concerned that our analytical sample might not be representative of the overall IID population. In Figure A1, we plot mean child income rank against parental income rank for different samples. Mobility patterns are quite robust to sample selection: we find that the rank-rank slope is 0.227 in the original IID population, 0.233 in a sample of children matched to the Census (with no restriction on whether their parents are matched or not), and 0.239 in our main analytical sample.

Variables Definitions The IID tax records include a variety of income measures for each person in a family. We use total individual income from all sources before tax as defined by the Canada Revenue Agency. This includes earnings, interest and investment income, self-employment net income, taxable capital gains/losses and dividends, and benefits. We use the Consumer Price Index to convert all dollar figures to 2016 Canadian dollars. To maximize sample size, we measure child income as the average pre-tax total income over 5 years between the ages of 27 and 31 inclusively.[3] Taking the average over 5 years for both

[1] See Corak and Heisz (1999) and Connolly et al. (2019) for a detailed description of the IID.

[2] Children matched to at least one Census wave represent 54% of all children in the IID. The second restriction further brings this fraction down to 31%. The matching procedure is described in the Online Data Appendix.

[3] Children born in 1985 – our most recent birth cohort – were 31 years old in 2016, the last year of tax data.

the parents and the child reduces the impact of transitory income shocks (Deutscher and Mazumder, 2020).

We define parental income as the average total family income (the sum of both parents' income) over 5 years, when the child is aged 15 to 19. In this sense, our measure of income reflects the monetary resources available to the child when he or she was a teenager, a period during which important educational investment decisions that may be affected by parental financial resources are made. Using parental income measured when the child is 10 to 19 years old, or when the parents are themselves 40 to 49 years old, or 45 to 54 years old, does not significantly affect our results.[4] We convert average income into percentile ranks for both parents and children. Percentiles are based on the national income distribution in the original, full IID population, and they are calculated separately for each children birth cohort.

From the Census, we obtain information on the occupation and field of study of children and their parents. There are 500 occupational codes, and 446 different fields of study. Individuals who filled out the Census but did not have an occupation because they were either not working at the time of the Census or had not worked the year prior to the Census are assigned to a "no occupation" category (about 6% of children in our sample). The field of study refers to the predominant discipline or area of study of the respondent's highest completed post-secondary certificate, diploma or degree. The fields of study are classified according to the Classification of Instructional Programs (CIP). Fields of study are only observed for individuals with some post-secondary schooling: 34% of children in our sample have no reported field of study, and are assigned to a "no field" category. Further details on the construction of the sample and the coding of variables are provided in the Online Data Appendix.

2 Intergenerational Income Mobility and Occupations

Patterns of Segregation across Occupations. For illustrative purposes, we begin our empirical investigation by plotting the distribution of parental income for selected children's occupations. We separately examine high-paying and low-paying occupations, and focus on occupations for which our sample sizes are large enough to allow for the estimation of the complete distribution of workers across parental income fiftiles. Figure 1 shows that these distributions are highly skewed for high-paying occupations such as physicians and lawyers.

[4]Appendix Figure A2 shows that rank-rank estimates of intergenerational income mobility are fairly robust to the age at which children's and/or parents' income is measured.

A staggering 12.4% of physicians grew up in families in the top 2% of the parental income distribution. In total, a majority (53%) of physicians had parents in the top quintile of the income distribution, confirming survey-based findings that medical students are disproportionately drawn from high-socioeconomic status backgrounds (Dhalla et al., 2002). Workers in low-paying occupations such as cooks and hairstylists are disproportionately drawn from low-income families, but the degree of concentration at the bottom of the distribution does not mirror the extreme degree of segregation that is observed in high-paying occupations.

To examine patterns of segregation across all 500 occupations, we then compute a Herfindahl-Hirschman concentration index for each occupation $HHI_o = \sum_p (s_{p|o})^2$, where $s_{p|o}$ is the share of children in occupation o whose parents are in the p-th percentile of the parental income distribution. If children in occupation o are uniformly distributed across parental income percentiles, then $HHI = 0.01$. If all children in occupation o were drawn from a single parental income percentile, then $HHI = 1$. In our sample, the average child is in an occupation with a HHI of 0.0111. The minimum value is 0.0101 and the maximum is 0.0457. Figure A3 plots the HHI against a measure of occupational returns (expressed in income percentile ranks).[5] The patterns confirm that concentration by parental income is asymmetric, with high-paying occupations being more segregated by parental income than low-paying ones, on average.

Occupational Returns and Income Mobility. We now estimate the fraction of the intergenerational income rank-rank relationship – our main measure of intergenerational income mobility – that can be accounted for by the occupations of children. To do so, we include children's occupation fixed effects in a regression of children income rank on a complete set of parental income rank indicators:

$$y_{io} = \sum_{p=1}^{100} \beta_p 1\{x_i = p\} + \sum_o \delta_o + \varepsilon_{io} \tag{1}$$

where y_{io} is the income rank of child i in occupation o and x_i is the income rank of their parents. We normalize the occupation fixed effects δ_o to have a mean of zero in the estimation sample.

The average income rank of children with parental income p, $\bar{y}_p = E[y_{io}|x_i = p]$, can be decomposed into a component reflecting within-occupation income transmission (β_p), and

[5] These returns are obtained via the estimation of equation (1), described below.

one capturing occupational segregation by parental income rank (Δ_p):

$$\bar{y}_p = \underbrace{\sum_o \left(\bar{y}_{o|p} - \hat{\delta}_o\right) s_{o|p}}_{\hat{\beta}_p} + \underbrace{\sum_o \hat{\delta}_o s_{o|p}}_{\Delta_p} \tag{2}$$

where $\bar{y}_{o|p}$ is the average income rank of children in occupation o belonging to group p, and $s_{o|p}$ is the share of children of group p who are in occupation o. This approach effectively corresponds to a regression-compatible Kitagawa-Oaxaca-Blinder-type decomposition of income gaps between 100 parental income groups (Fortin, 2008). Equation (2) highlights how the contribution of occupations (the "explained" component Δ_p) depends on the distribution of estimated occupational returns $\hat{\delta}_o$, as well as on the distribution of children's occupations across parental income groups $s_{o|p}$.[6]

Figure 2, panel A plots average income ranks $\bar{y}_p$ of children as a function of parental income ranks – the unconditional rank-rank relationship (small circles). The series in triangles plots the estimated β_p's, the conditional rank-rank relationship. The vertical distance between the two series is equal to Δ_p, the occupational segregation term. As a summary measure, we report the corresponding linear slopes, which are equal to 0.239 and 0.170, respectively. This implies that segregation into occupations by parental income accounts for close to a third (0.069/0.239 = 29%) of the intergenerational transmission of income.[7] These linear slope coefficients mask important non-linearities at the tails. The rank-rank relationship is considerably steeper below the 10th percentile, and the top percentile is a clear outlier. The non-linearities are equally apparent in the conditional series, suggesting they are mostly unrelated to occupational choices. Children from top 1-percent families obtain particularly high incomes, but little of it is explained (in a statistical sense) by the occupations they hold.

Figure 2, panel B unpacks the difference between the unconditional and conditional slopes by plotting the estimated return of each occupation $\hat{\delta}_o$ against the average parental income within each occupation.[8] The size of the dots indicates the proportion of children in each occupation. For illustrative purposes, we highlight in red selected low-mobility occupations – those that have both high (low) average parental income and high (low) economic returns and

[6]Sampling error may lead us to slightly overstate the variance of occupation fixed effects, and therefore the magnitude of the occupational segregation terms. Yet, with over 5,000 observations per occupation on average, this bias is likely minimal.

[7]Appendix Figure A2 shows that these results are robust to alternative measures of income.

[8]Note that re-scaling the linear fit slope of the plotted relationship (1.11) by the between-occupation share of the variance of parental income $\frac{Var(\bar{x}_o)}{Var(x_i)} = 0.062$ returns the difference between unconditional and conditional rank-rank slopes of 0.069.

thereby contribute to intergenerational income persistence – as well as selected high-mobility occupations in green.[9] For example, the average specialist physician had parents at the 74^{th} percentile of the income distribution, and benefits from an occupational return of +23 income percentiles. In contrast, the average cook grew up in a family at the 44^{th} percentile of the parental income distribution, 30 percentiles below specialist physicians, and obtains a negative occupational return of -15 income percentiles, 38 percentile ranks below specialist physicians. Occupations that foster relative mobility (in an accounting sense) are found in the northwest and southeast quadrants. Several occupations related to the mining, oil and gas and construction industries have occupational returns similar to that of dentists and physicians (between +20 and +30 percentile ranks), but workers in these occupations have much lower average parental income. On the other hand, children in occupations associated with the arts (e.g. actors, dancers and musicians) have above-average parental income but large negative returns of about -20 percentile ranks.

Heterogeneity by Gender. In Figure 3, we present unconditional and conditional rank-rank slopes estimated separately by child gender. Despite important differences in levels, the unconditional slopes are very similar for men and women. Overall, occupations have greater explanatory power for women. The slope of the linear fit of the relationship between the segregation term Δ_p and parental income rank p is 0.085 for women, and 0.07 for men. This is in part because high upward mobility occupations tend to be concentrated in the mining and construction industries, sectors in which women are underrepresented. In fact, while 8.7% of men are in occupations in the top-left quadrant (i.e. positive occupational returns and average parental income below the median), only 0.08% of women are.[10]

While men's and women's occupational returns are strongly correlated (0.91), there are some important differences. For instance, women obtain substantially greater returns to pursuing engineering occupations, as well as to pursuing careers in medicine and pharmacy. To quantify the importance of these differences in occupational returns, we calculate a counterfactual linear slope of the segregation term as a function of parental income rank for women by assigning them men's returns, holding the distribution of occupations constant. The resulting value is 0.063, which is smaller than the actual number for men.[11] Hence, the

[9]The complete set of estimates shown on these figures is available in the paper's replication package.

[10]This finding is due to under-representation of women in these occupations rather than to gender differences in the returns to these occupations. If we assign men's occupational returns to women holding the distribution of occupations constant, the fraction of women in the top-left quadrant drops to 0.07%.

[11]Correspondingly, assigning women's occupational returns to men produces a counterfactual slope of the segregation term for men of 0.094.

greater dispersion of occupational returns for women contributes significantly to – and even over-explains – gender differences in intergenerational mobility.

Fields of Study. In this subsection, we examine whether occupational sorting can be traced back to earlier educational choices. While some occupations require formal degrees in specific fields of study, many do not. That is, the mapping between fields of study and occupations is not deterministic and may vary across parental income groups.

Figure A4, panel A shows that about half of the occupational segregation gradient – the linear slope of Δ_p as a function of parental income rank p – is accounted for by differences in the fields of study children pursue across parental income groups.[12] This implies that even conditional on fields of study, children of rich households are systematically more likely to work in high-paying occupations. In panel B, we show that fields of study have overall fairly limited explanatory power for intergenerational income mobility. The income rank-rank slope conditional on field of study fixed effects is 0.188, a difference of 0.051 relative to the unconditional rank-rank slope of 0.239.[13]

3 Occupational Persistence Across Generations

Why do children from different parental income groups end up in different occupations? One natural explanation is that parental occupations play a key role in determining children's career choices. In this section, we unpack patterns of occupational transmission across generations and relate those to differences in the distribution of children's occupations across parental income groups.

Patterns of Occupational Following. As a first step, we document whether children are more likely to pursue a given occupation if their parents held that specific occupation or a similar one. For each occupation, we estimate the following regression:

$$O_{ioq} = \pi_0^o + \pi_1^o \mathbf{1}\{q = o\} + \pi_2^o \mathbf{1}\{q \in B_o, q \neq o\} + \gamma_{t(i)}^o + \varepsilon_{ioq} \tag{3}$$

[12]This result is obtained by assigning each child the value of $\hat{\delta}_o$ for their occupation, and regressing these occupational returns on parental income rank dummies and field of study fixed effects in the micro data.

[13]Since fields of study are only observed for individuals with some post-secondary schooling, they also reflect differences in educational attainment. This likely explains why the explanatory power of fields of study appears greater at the top of the parental income distribution.

where O_{ioq} is an indicator variable taking a value of 1 if child i works in occupation o and zero otherwise, $\gamma^o_{t(i)}$ is a vector of birth cohort fixed effects, $\mathbf{1}\{q = o\}$ indicates whether either parent of child i works in occupation o, and $\mathbf{1}\{q \in B_o, q \neq o\}$ indicates whether either parent works in an occupation (other than occupation o) that belongs to the same broad category B_o as occupation o. We use 3-digit occupation codes (of which there are 141) to operationalize these broader categories. In several cases, there are only 2 or 3 occupations in the category, so the coefficient π^o_2 represents the increase in probability of pursuing occupation o that is associated with one's parent working in the occupation that is "closest" to occupation o.

Figure 4 presents estimates of π^o_1 and π^o_2 in panels A and B, respectively.[14] In both cases, the vast majority of coefficients are positive, indicating that a child is more likely to pursue an occupation if either of their parents is working in that same occupation or a similar one. On average, having a parent in a given occupation increases the probability of pursuing that same occupation by 2.1 percentage points. Relative to children who do not have a parent in a related occupation, this represents a 9-fold increase on average. Occupational following is fairly "local": on average, having a parent in an occupation similar to occupation o increases the probability of pursuing occupation o by only 0.5 percentage point. This highlights the importance of using detailed occupational codes for unpacking patterns of occupational transmission.

There is substantial heterogeneity in intergenerational persistence across occupations. For ease of exposition, in Figure 4 we highlight selected occupations for which the persistence coefficients are strikingly high. For example, the probability of being a lawyer is 6.2 percentage points greater if one has a lawyer parent, or 15 times higher than for children without parents in related occupations. Occupational following is particularly strong in construction trades and agricultural occupations. For example, having a parent working as a general farm worker increases the probability of working in that occupation by about 8.5 percentage points, a 18-fold increase relative to children without a parent that occupation or a similar one.

Occupations differ considerably in terms of size (i.e. number of workers), making comparisons of conditional probabilities hard to interpret. To address this, we construct a fluidity index that we define as the ratio of the probability of picking occupation o conditional on *not* having a parent in that or a similar occupation ("outsiders"),[15] and the unconditional

[14] Appendix Figure A5 reproduces these results for paternal and maternal occupations, separately. The patterns are fairly similar across parental gender.

[15] This probability is measured as the average fitted values from eq. (3), excluding the contribution of the coefficients π^o_1 and π^o_2.

probability of pursuing occupation o ($\bar{O}_o$). The index takes a value of one if outsiders – children without a parent in that or a similar occupation – are just as likely to pick that occupation as those who have a parent in that or a similar occupation (i.e., if $\pi_1^o = \pi_2^o = 0$). It takes a value of zero if only children of parents in that or a similar occupation ever pursue the occupation.

Estimates are presented in Figure 4, panel C. The average fluidity index is 0.97, suggesting a very high degree of occupational mobility. There is a positive relationship between average parental income and fluidity. Occupations for which children come disproportionately from low-income families are generally the most persistent across generations. That is, while children of high-income families tend to pursue high-paying occupations, little of this is explained by direct occupational transmission. If anything, occupational persistence is greater among occupations in which low-income families are over-represented, such as farm workers and fishermen/women. Laband and Lentz (1983) argue these are occupations where the transmission of knowledge and the inheritance family businesses may be important.

Occupational Transmission and Income Mobility. Next, we quantify the fraction of the occupational segregation term (Δ_p) that is explained by differences in parental occupations. First, let $\nu_{q|p}$ denote the share of parents in income group p who work in occupation q and let $s_{o|q,p}$ be the share of children of parents in occupation q and group p who work in occupation o. For any parental income group p, the values of $s_{o|q,p}$ therefore provide a complete, *income group-specific* 500×500 occupational transition matrix between the two generations, for a total of 100 such matrices. Then, the share of children in occupation o for each parental income rank p is $s_{o|p} = \sum_q s_{o|q,p}\nu_{q|p}$, which emphasizes that children from different income groups possibly transition into different occupations because (1) their parents belong to different occupations (i.e., differences in starting points $\nu_{q|p}$) and (2) they have different transition paths $s_{o|q,p}$ (i.e. the same starting points lead to different occupations for children).

To assess the role of occupational persistence, we break down the segregation component as follows:

$$\Delta_p = \underbrace{\sum_o \hat{\delta}_o \sum_q \left(s_{o|q}\nu_{q|p}\right)}_{\text{Differences in parental occupations}} + \sum_o \hat{\delta}_o \sum_q \left(s_{o|q,p} - s_{o|q}\right)\nu_{q|p} \tag{4}$$

where $s_{o|q}$ is the share of children of parents in occupation q, in the full population, who work in occupation o. We focus on calculating the first term, which quantifies the role of having different occupational "origins" ($\nu_{q|p}$) by restricting the occupational transition matrix ($s_{o|q}$) to be the same for all parental income groups. That is, counterfactual distributions of children occupations are obtained by replacing the parental income-specific transition weights ($s_{o|q,p}$) with population-wide weights ($s_{o|q}$). We conduct this exercise twice – once using the fathers' occupations to construct the transition matrix, and once using the mothers'.

Figure 5 shows counterfactual occupational segregation terms under the scenario where occupations are passed-on the same way across parental income groups, but children from different groups still differ in terms of the occupations they inherit. Throughout the parental income distribution, differences in paternal and maternal occupations account only for a small fraction of overall occupational segregation. The linear slope of the counterfactual segregation component based on paternal occupations is 0.022, less than half that of the actual occupational segregation slope of 0.069, and equivalent to 10% of the overall income rank-rank slope of 0.239. The explanatory power of mothers' occupations is even smaller.[16] Hence, most of the variation in average occupational returns across parental income groups is due to the fact that children from high- and low-income families have different transition paths. Put differently, children of high-income families cluster in high-paying occupations largely independently of their parents' occupations. In Appendix Figure A6 we reproduce this analysis separately by child gender. There are some asymmetries, with mothers' occupations contributing relatively more to daughters' occupational segregation than sons'.

Differential occupational returns The previous analysis quantifies the contribution of occupational following to differences in the occupations children pursue, but ignores the possibility that parental occupations may affect the economic returns children obtain from working in different occupations. For example, children of lawyers may not only be more likely to become lawyers, but may also benefit from their parents' network relative to other lawyers who do not have lawyer parents (Raitano and Vona, 2021), thereby increasing their economic returns to practicing law. Alternatively, nepotism and barriers to entry may allow less-talented children with lawyer parents to enter the profession, and children who overcome these barriers despite not having lawyer parents plausibly have relatively higher earnings potential. In such a case, one may instead expect occupational followers to have lower occupational returns on average.

Intergenerational transmission of occupation-specific human capital and the prevalence of occupation-specific nepotism could both contribute to within-occupation income differences

[16] Even if paternal and maternal occupations were perfectly independent of each other, the combined (additive) explanatory power of the two would still be below half of the occupational segregation slope.

between children with and without parents in that occupation. Such heterogeneous returns may, in turn, partly explain why parental income continue to predict child income even conditional on occupational attainment.

To examine whether this is the case, we augment equation (1) with a full set of interactions between children's occupation fixed effects and an indicator for having a parent working in the same broad occupational category, where broad categories correspond to 3-digit occupational groupings as before. Figure A7, panel A shows how occupational returns (conditional on parental income) for children with a parent in the same occupational category – occupational followers – compare to those for children without a parent in that occupational category. The two are strongly correlated (0.96). On average, occupational returns are 0.7 percentile points greater for occupational followers, but there is no evidence that this premium is systematically greater in high-paying occupations. As a result, heterogeneous occupational returns by parental occupation are virtually inconsequential for overall income mobility. Panel B shows that the conditional rank-rank relationship that allows for heterogeneous occupational returns is essentially identical to that for constant returns.

4 Conclusion

We document systematic segregation by parental income into occupations and fields of study, particularly so for high-paying occupations. Yet, the occupations that children pursue account only for a third of the intergenerational transmission of income. In addition, most of this segregation into occupations cannot be traced back to the occupations of parents. Rather, the fact that children of high-income families disproportionately pursue high-paying occupations is mostly independent of the parental occupations they inherited. In contrast, occupational following is relatively more prevalent in low-paying occupations.

Overall, this paper suggests that occupations and fields of study play a somewhat limited role as drivers of intergenerational income mobility. One implication is that patterns of occupational persistence across generations significantly understate the degree of persistence in income. Our findings also highlight the importance of considering alternative, post-schooling factors that may contribute to the correlation between parental income and within-occupation child residual income. For instance, Corak and Piraino (2011) show that about 40% of men have, at some point, worked for an employer that also employed their father at some time. More generally, firms play an important role for earnings inequality (Card et al., 2018; Song et al., 2019; Lamadon et al., 2022). We suspect that parental income may be associated with the type of firms children end up working for, even within occupations, if parental networking is important (Kramarz and Skans, 2014; San, 2021; Staiger, 2021). We leave this investigation for future work.

References

Aina, Carmen and Cheti Nicoletti, "The intergenerational transmission of liberal professions," *Labour Economics*, 2018, *51*, 108–120.

Bello, Salvatore Lo and Iacopo Morchio, "Like father, like son: Occupational choice, intergenerational persistence and misallocation," *Quantitative Economics*, 2022, *13* (2), 629–679.

Björklund, Anders, Markus Jäntti et al., "Intergenerational mobility of socio-economic status in comparative perspective," *Nordic Journal of Political Economy*, 2000, *26*, 3–32.

Blanden, Jo, "Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology," *Journal of Economic Surveys*, 2013, *27* (1), 38–73.

Blau, Peter M and Otis Dudley Duncan, "American occupational structure," 1967.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

__ , __ , Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Connolly, Marie, Catherine Haeck, and Jean-William Laliberté, "Parental Education and the Rising Transmission of Income between Generations," in Raj Chetty, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell, eds., *Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth*, NBER, 2022, chapter 2.

__ , Miles Corak, and Catherine Haeck, "Intergenerational Mobility between and within Canada and the United States," *Journal of Labor Economics*, 2019, *37* (S2), S595–S641.

Corak, Miles and Andrew Heisz, "The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data," *Journal of Human Resources*, 1999, *34* (3), 504–533.

__ and Patrizio Piraino, "The intergenerational transmission of employers," *Journal of Labor Economics*, 2011, *29* (1), 37–68.

Dahl, Molly and Thomas DeLeire, *The association between children's earnings and fathers' lifetime earnings: estimates using administrative data* 2008.

Deutscher, Nathan and Bhashkar Mazumder, "Intergenerational mobility across Australia and the stability of regional estimates," *Labour Economics*, 2020, *66*, 101861.

Dhalla, Irfan A, Jeff C Kwong, David L Streiner, Ralph E Baddour, Andrea E Waddell, and Ian L Johnson, "Characteristics of first-year students in Canadian medical schools," *Cmaj*, 2002, *166* (8), 1029–1035.

Doepke, Matthias and Fabrizio Zilibotti, "Parenting with style: Altruism and paternalism in intergenerational preference transmission," *Econometrica*, 2017, *85* (5), 1331–1371.

Erikson, Robert, John H Goldthorpe, and Lucienne Portocarero, "Intergenerational class mobility in three Western European societies: England, France and Sweden," *The British Journal of Sociology*, 1979, *30* (4), 415–441.

Fortin, Nicole M, "The gender wage gap among young adults in the united states the importance of money versus people," *Journal of Human Resources*, 2008, *43* (4), 884–918.

Jácome, Elisa, Ilyana Kuziemko, and Suresh Naidu, "Mobility for All: Representative Intergenerational Mobility Estimates over the 20th Century," Technical Report, National Bureau of Economic Research 2021.

Kramarz, Francis and Oskar Nordström Skans, "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, 2014, *81* (3), 1164–1200.

Laband, David N and Bernard F Lentz, "Like father, like son: Toward an economic theory of occupational following," *Southern Economic Journal*, 1983, pp. 474–493.

__ and __ , "Self-recruitment in the legal profession," *Journal of Labor Economics*, 1992, *10* (2), 182–201.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 2022, *112* (1), 169–212.

Lentz, Bernard F and David N Laband, "Why so many children of doctors become doctors: Nepotism vs. human capital transfers," *Journal of Human Resources*, 1989, pp. 396–413.

Long, Jason and Joseph Ferrie, "Intergenerational occupational mobility in Great Britain and the United States since 1850," *American Economic Review*, 2013, *103* (4), 1109–37.

Mazumder, Bhashkar and Miguel Acosta, "Using occupation to measure intergenerational mobility," *The ANNALS of the American Academy of Political and Social Science*, 2015, *657* (1), 174–193.

Mocetti, Sauro, "Dynasties in professions and the role of rents and regulation: Evidence from Italian pharmacies," *Journal of Public Economics*, 2016, *133*, 1–10.

__ , Giacomo Roma, and Enrico Rubolino, "Knocking on parents' doors: regulation and intergenerational mobility," *Journal of Human Resources*, 2020, pp. 0219–10074R2.

Modalsli, Jørgen, "Intergenerational mobility in Norway, 1865-2011," *The Scandinavian Journal of Economics*, 2017, *119* (1), 34–71.

Raitano, Michele and Francesco Vona, "Nepotism vs. specific skills: the effect of professional liberalization on returns to parental background of Italian lawyers," *Journal of Economic Behavior & Organization*, 2021, *184*, 489–505.

San, Shmuel, "Who Works Where and Why? Parental Networks and the Labor Market," 2021.

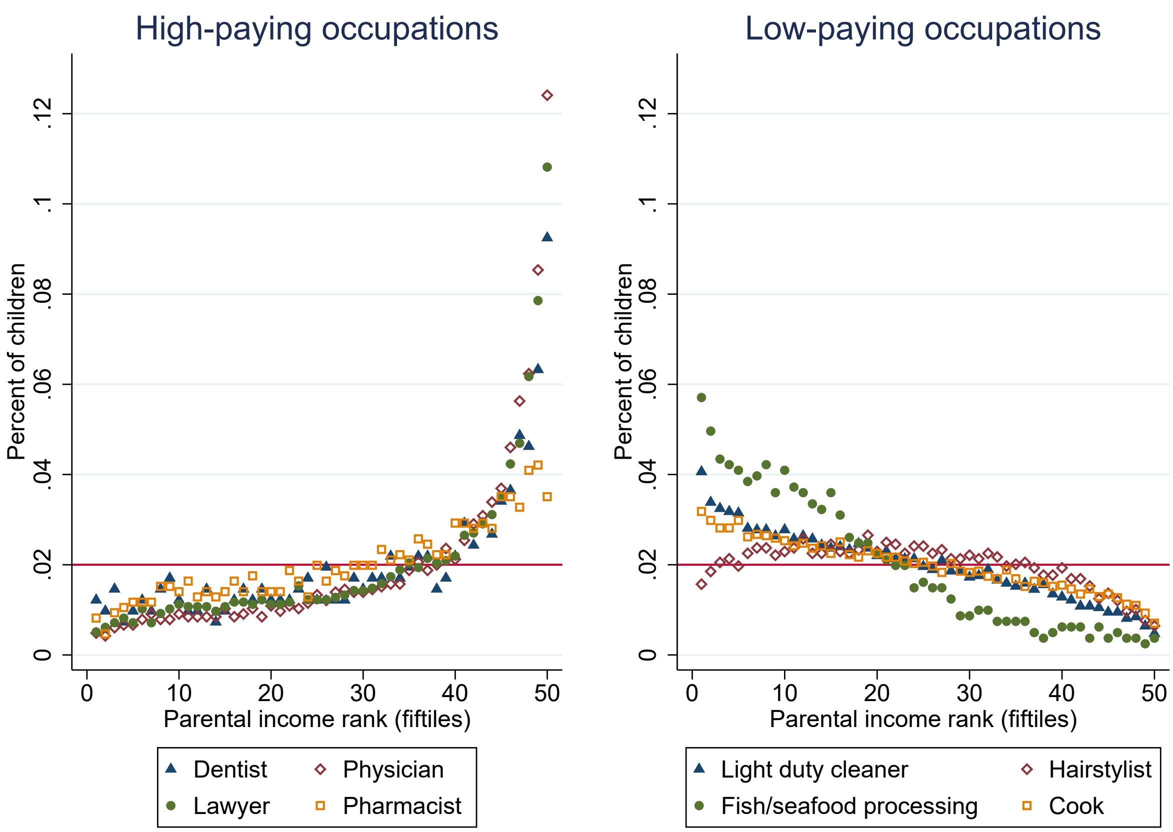
Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly Journal of Economics*, 2019, *134* (1), 1–50.

Staiger, Matthew, "The Intergenerational Transmission of Employers and the Earnings of Young Workers," 2021.

Torche, Florencia, "Analyses of intergenerational mobility: An interdisciplinary review," *The ANNALS of the American Academy of Political and Social Science*, 2015, *657* (1), 37–62.

Tables and Figures

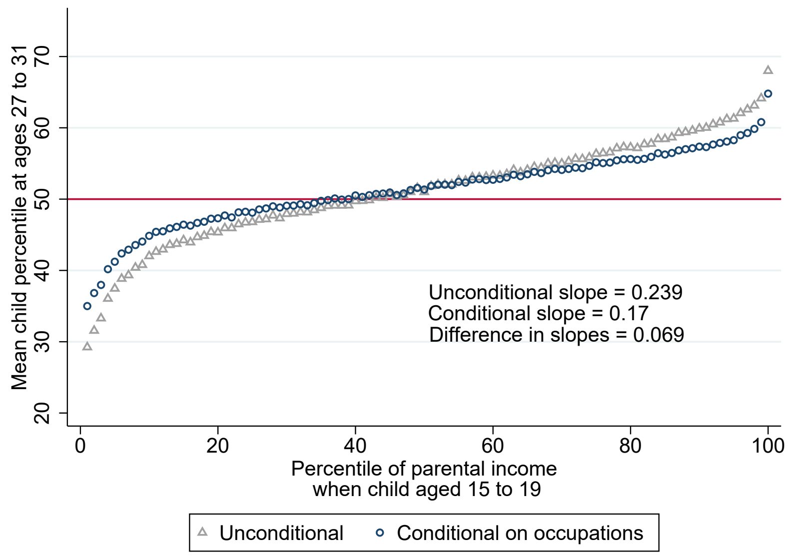
Figure 1: Distribution of Parental Income for Selected Occupations of Children



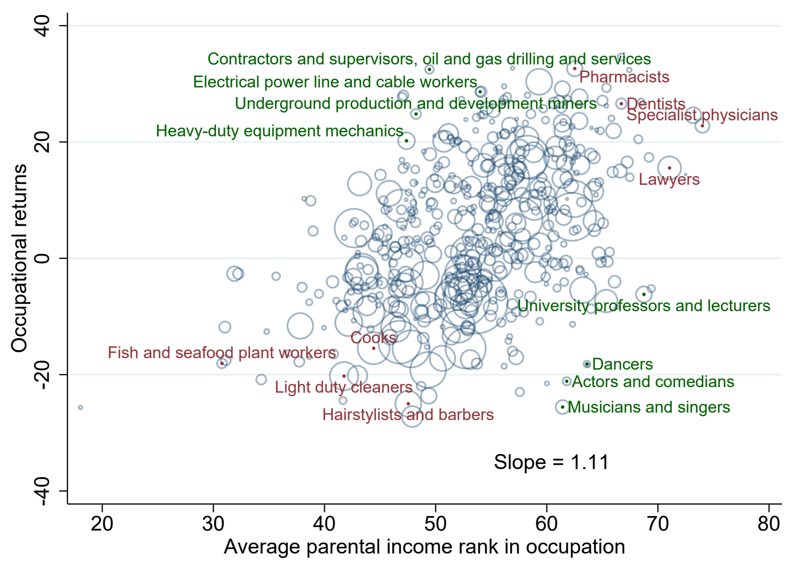
Notes: This figure shows the distribution of children across parental income fiftiles, separately for selected occupations (the children's occupation). The horizontal red bar at 0.02 indicates a uniform parental income distribution for reference.

Figure 2: Intergenerational Income Mobility and Children's Occupations

Panel A: Child-Parent Income Rank-rank Relationship



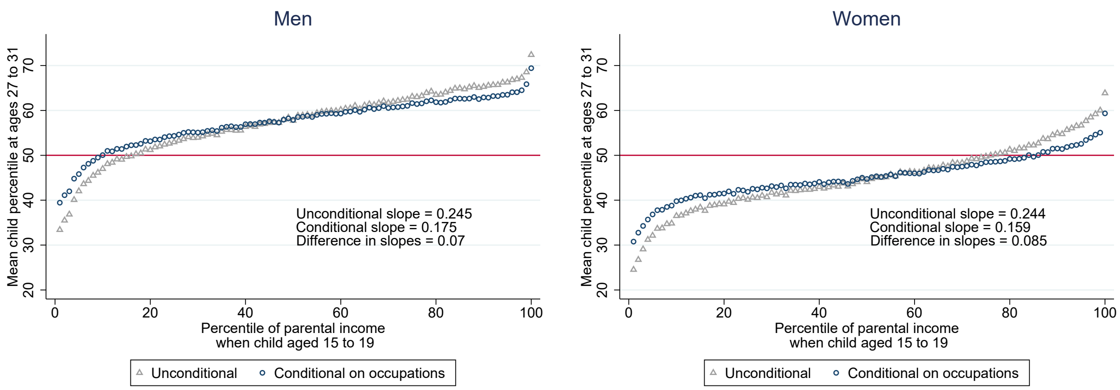
Panel B: Occupational Returns and Average Parental Income



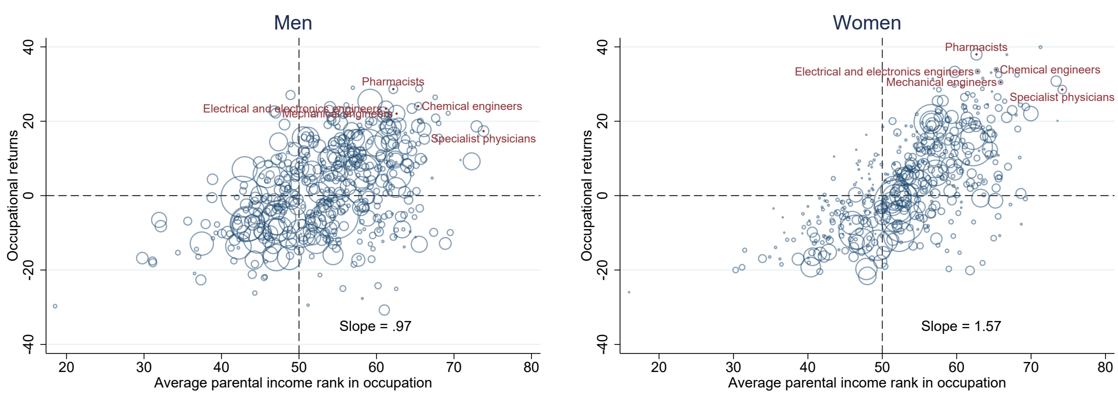
Notes: Panel A shows mean child income percentiles for each parental income decile. Grey triangles show unconditional means, whereas blue circles indicate conditional means accounting for differences in occupations, as per equation 1. The difference in linear slopes represents the contribution of children's occupations to income mobility. Panel B plots the estimates occupational returns from equation 1 against average parental income rank in each occupation. We highlight in red some occupations that contributes to the persistence of income across generations, and in green occupations that foster income mobility. The category "no occupation", which represents about 6% of the sample, is omitted from the scatter plot for visual clarity.

Figure 3: Intergenerational Income Mobility and Children's Occupations, by Gender

Panel A: Child-Parent Income Rank-rank Relationship



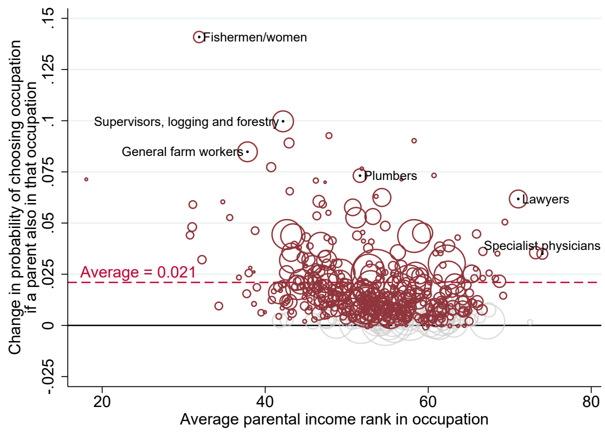
Panel B: Occupational Returns and Average Parental Income



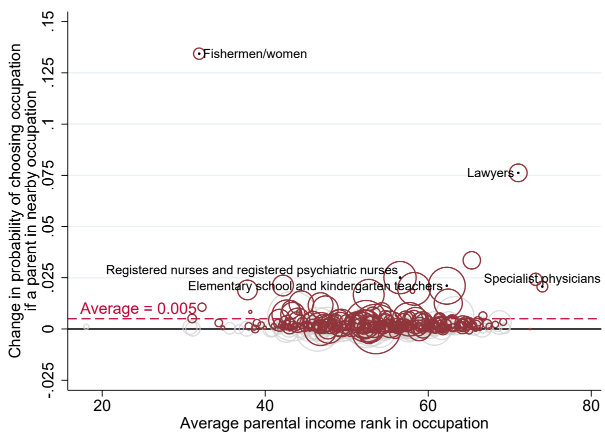
Notes: This figure reproduces Figure 2, separately by the child's gender. Estimation of equation (1) is done separately by gender, so that occupational returns are gender-specific.

Figure 4: Occupational Following Across Generations

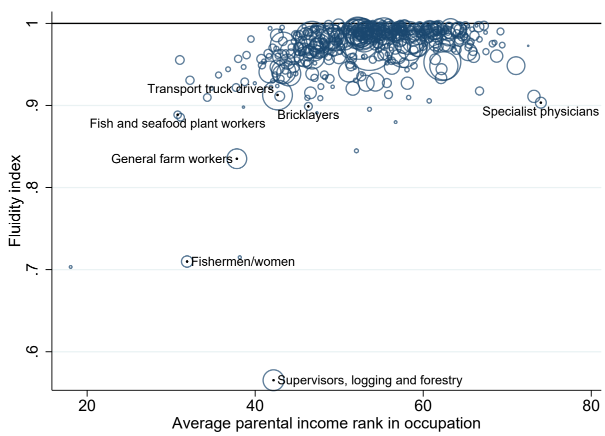
Panel A: Change in probability of pursuing an occupation if at least one parent is in that occupation



Panel B: Change in probability of pursuing an occupation if at least one parent is in similar (but not same) occupation



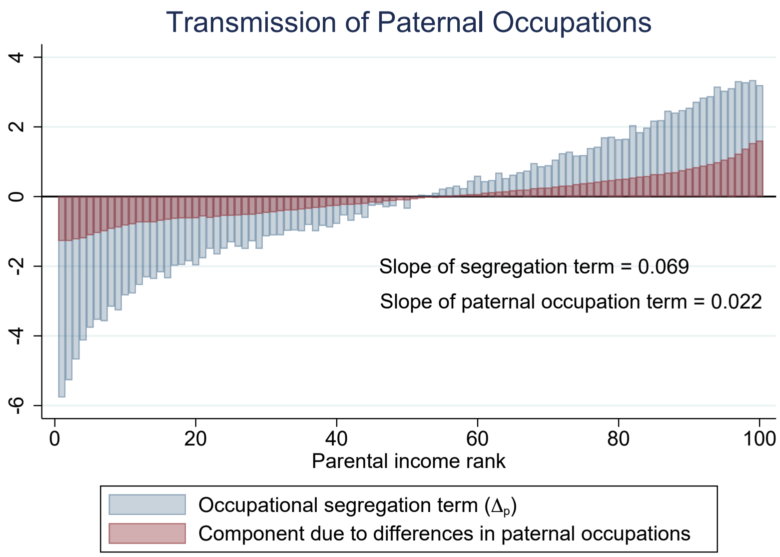
Panel C: Fluidity index by occupation



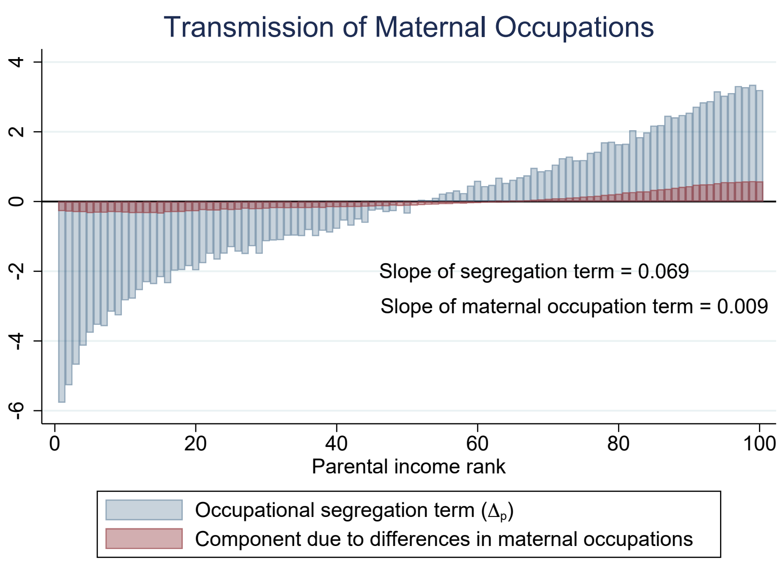
Notes: Panels A and B show estimates of π_1^o and π_2^o from equation (3). Regression coefficients that are statistically significant at the 95% level are depicted in maroon, and those that are not statistically significant at that level are shown in light grey. The category "no occupation" is omitted for visual clarity. The y-axis is trimmed at 0.15 for visual clarity. This results in dropping about 2 or 3 outlier occupations (out of 500) per figure. Panel C shows fluidity indices for all 500 occupations, again omitting the "no occupation" category.

Figure 5: Intergenerational Income Mobility and the Transmission of Occupations

Panel A: Transmission of Paternal Occupations



Panel B: Transmission of Maternal Occupations

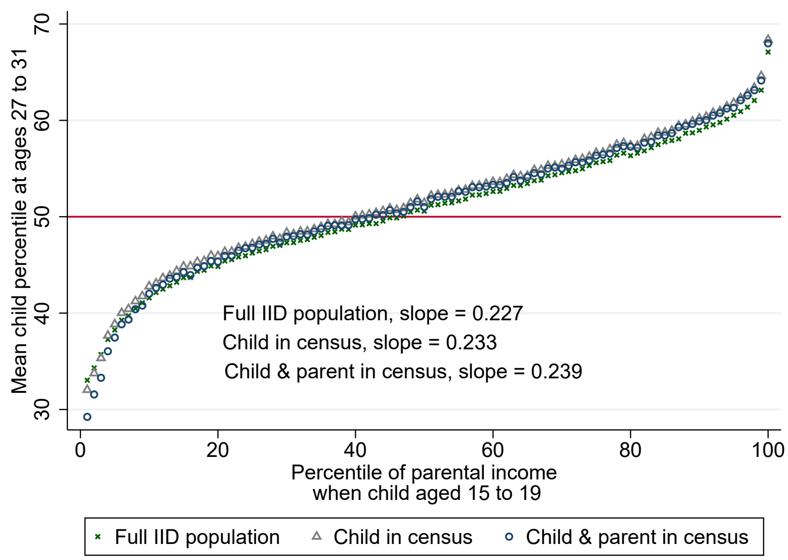


Notes: This figure presents estimates of counterfactual mean income ranks under the scenario in which the occupational transition matrix is the same for all income groups. Blue bars show the total contribution of occupations to income mobility, and red bars show the counterfactual contribution that isolates the role of the transmission of parental occupations. Panel A is based on differences in paternal occupations alone, whereas panel B isolates differences due to maternal occupations alone.

Appendix for Online Publication

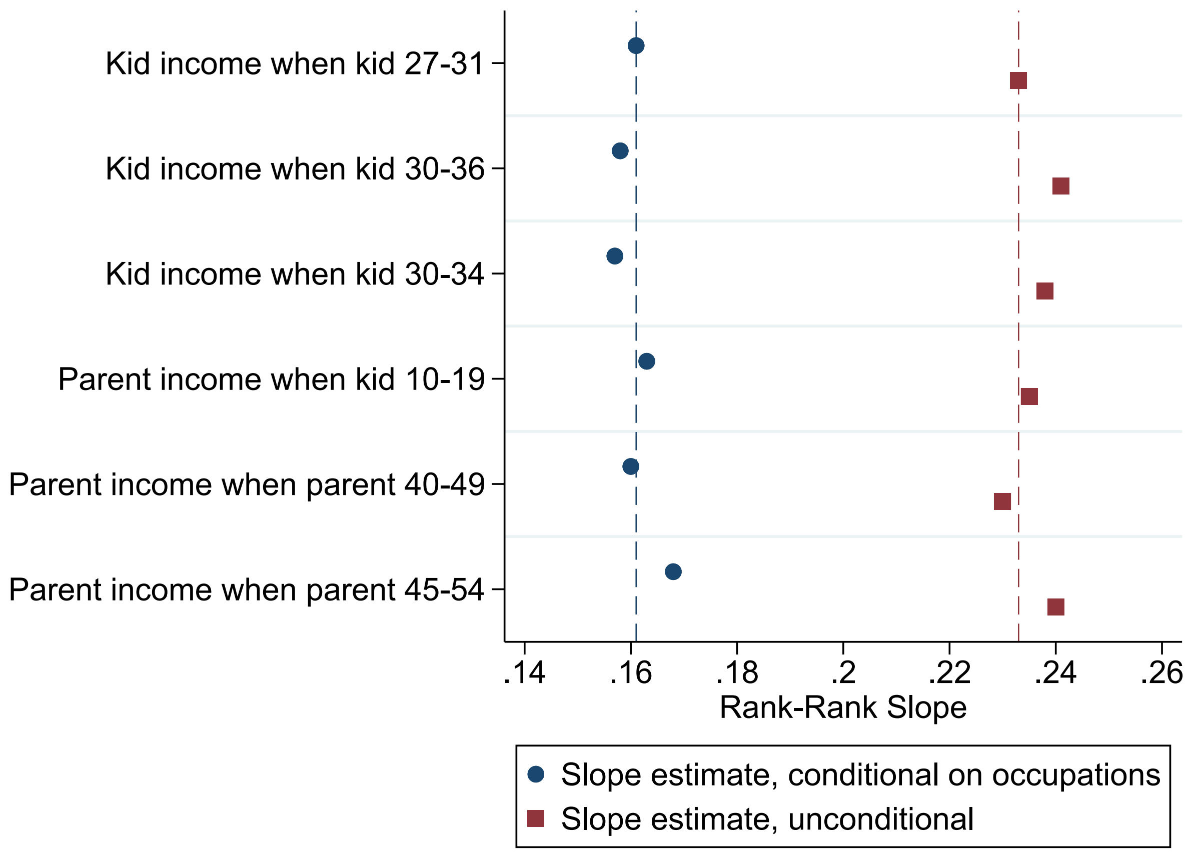
Appendix Figures

Figure A1: Intergenerational Income Mobility, Sensitivity to Sample Selection



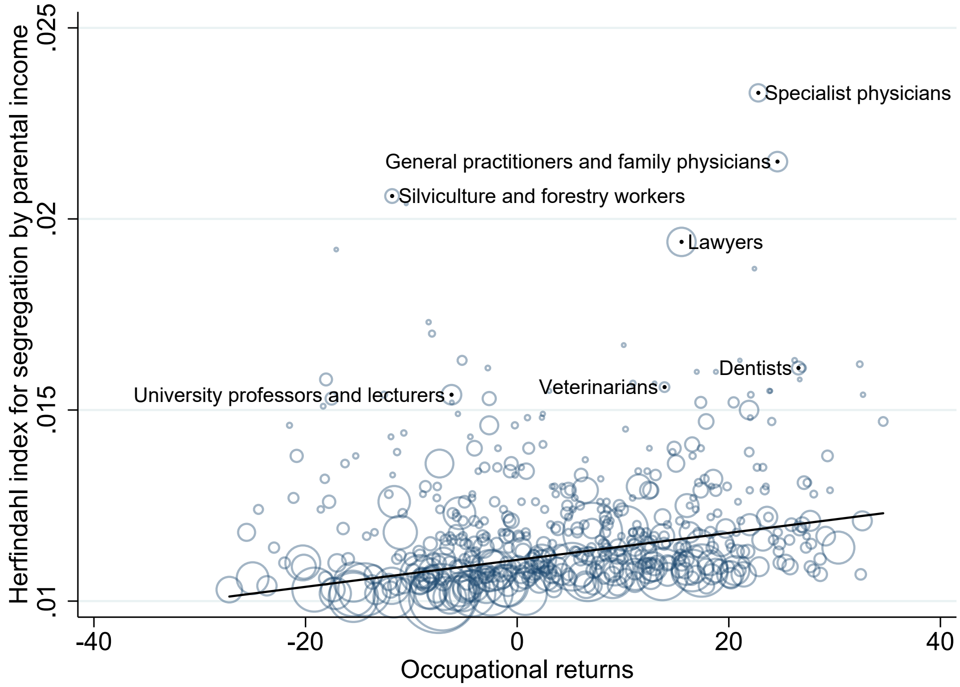
Notes: This figures shows mean child percentile ranks by parental income ranks for three different samples. The green x's show estimates for the entire IID sample. The grey triangles restrict the sample to children who are matched to at least one wave of the Census. The blue circles further restrict the sample to children who have at least one parent matched to at least one wave of the Census.

Figure A2: Intergenerational Income Mobility, Robustness



Notes: This figure shows linear slope estimates of the unconditional income rank-rank relationship, and the corresponding relationship conditional on children's occupations. In all specifications, the sample consists of all children matched to at least one wave of the Census. Across rows, we consider different measures of children and parental income ranks. The vertical lines indicate our baseline estimates.

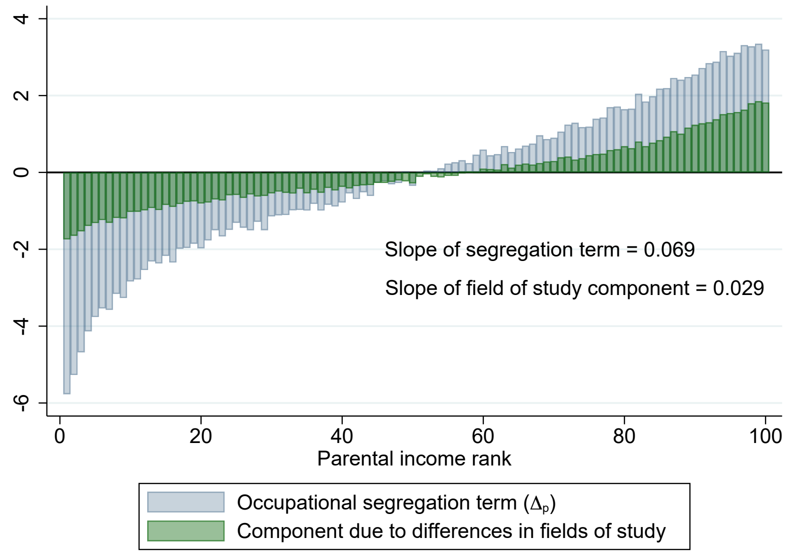
Figure A3: Segregation into Occupations by Parental Income, Herfindahl-Hirschman Index



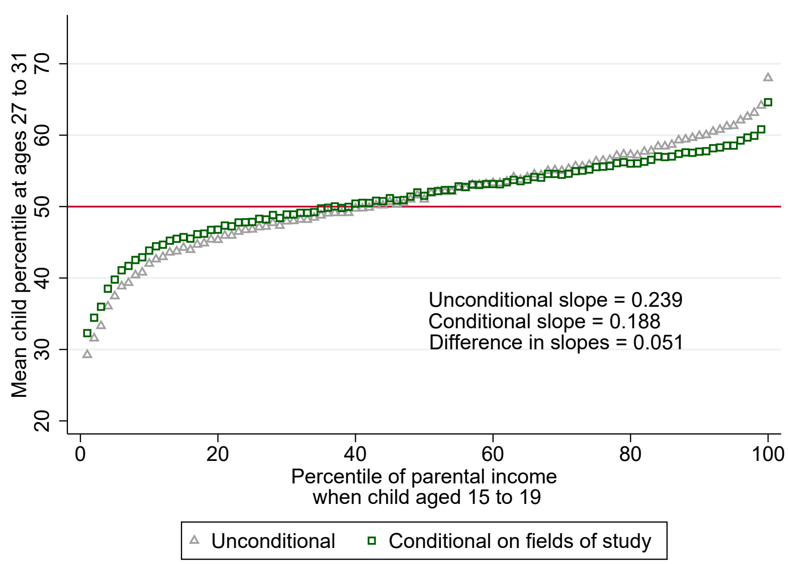
Notes: This figures plots Herfindahl-Hirschman segregation indices against estimated occupational returns. Each dot represents one occupation. The black line shows the linear fit. The vertical axis is truncated at 0.025 for visual clarity, which results in dropping 3 occupations.

Figure A4: Intergenerational Income Mobility and Children's Fields of Study

Panel A: Occupational segregation and Fields of Study



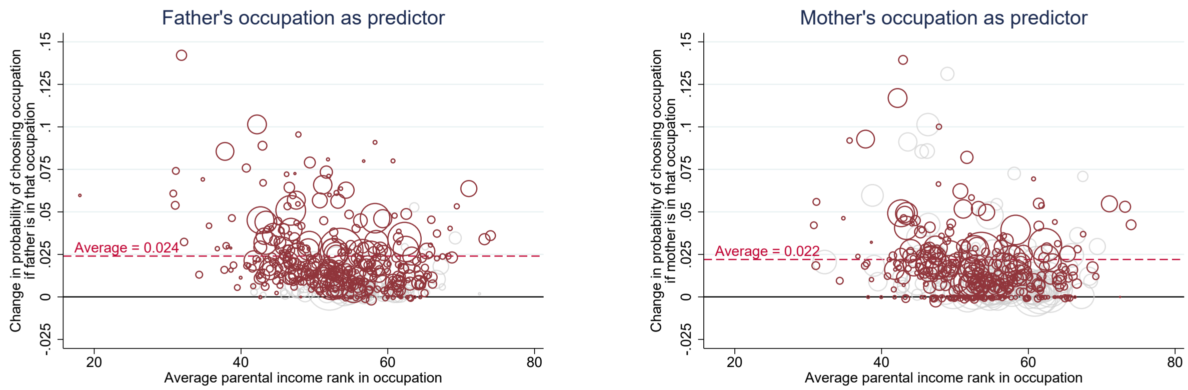
Panel B: Child-Parent Income Rank-rank Relationship



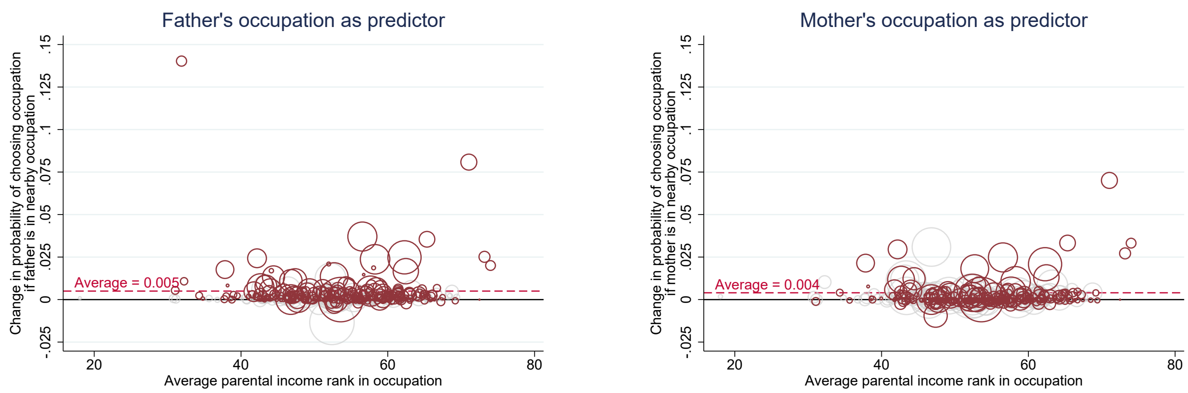
Notes: Panel A shows the fraction of occupational segregation by parental income that can be traced back to differences in fields of study. Blue bars show the total contribution of occupations to income mobility, and green bars show the component that is explained by fields of study. To calculate that component, we regress $\hat{\delta}_o$ on parent income rank dummies as well as a full set of field of study fixed effects. The estimated parental income rank coefficients represent the fraction of occupational segregation that operates within fields of study. The green bars are the difference between the total occupational segregation Δ_p and the fraction that operates within fields of study. Panel B replicates Figure 2, substituting field of study fixed effects for occupation fixed effects in equation (1).

Figure A5: Occupational Following Across Generations, Separately for Paternal and Maternal Occupations

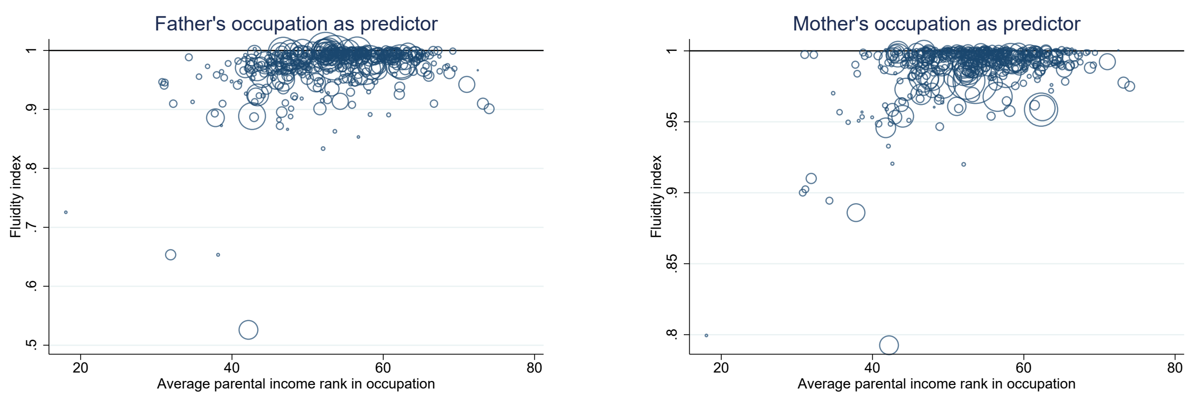
Panel A: Change in probability of pursuing an occupation if parent is in that occupation



Panel B: Change in probability of pursuing an occupation if parent is in similar (but not same) occupation

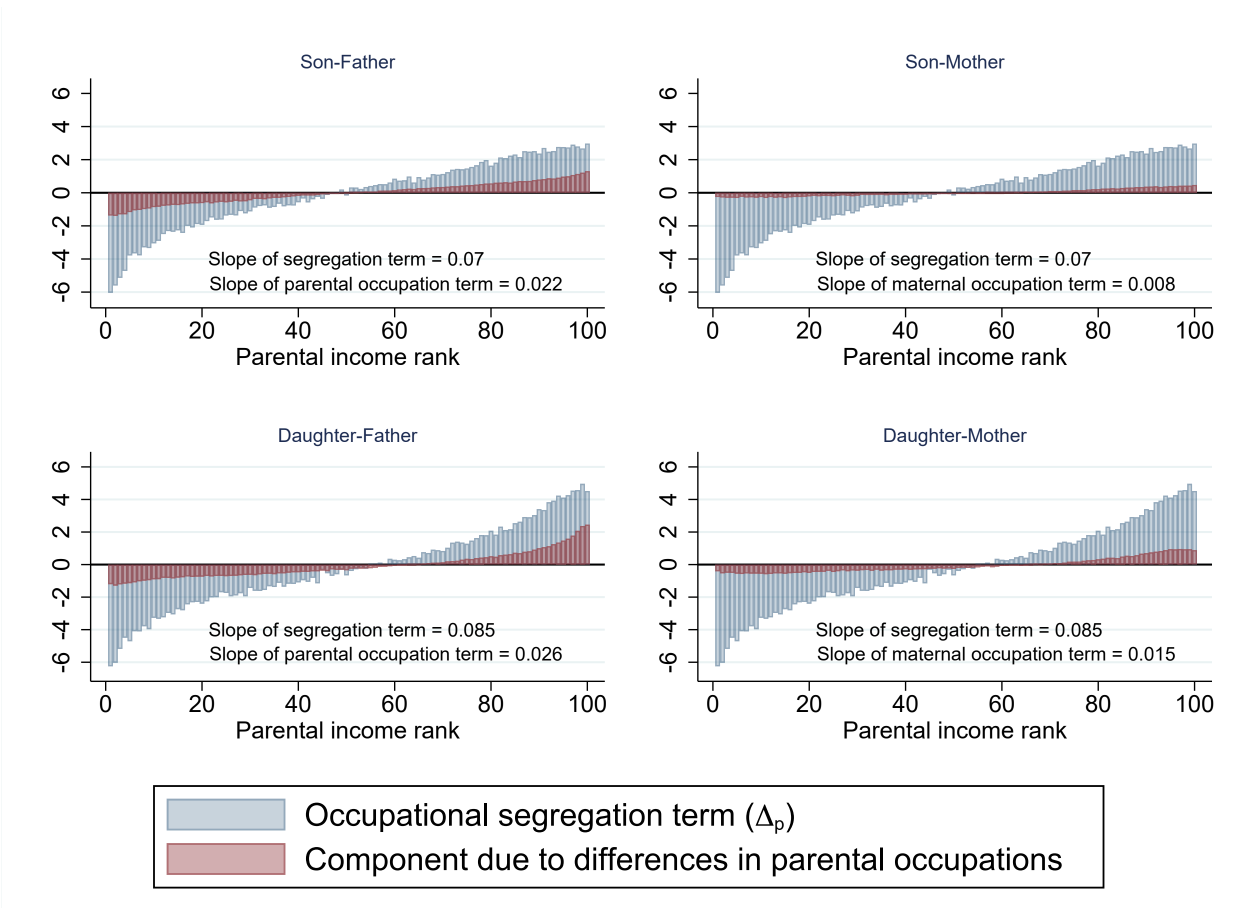


Panel C: Fluidity index by occupation



Notes: This figure reproduces Figure 3, using the father's and the mother's occupation as separate predictors.

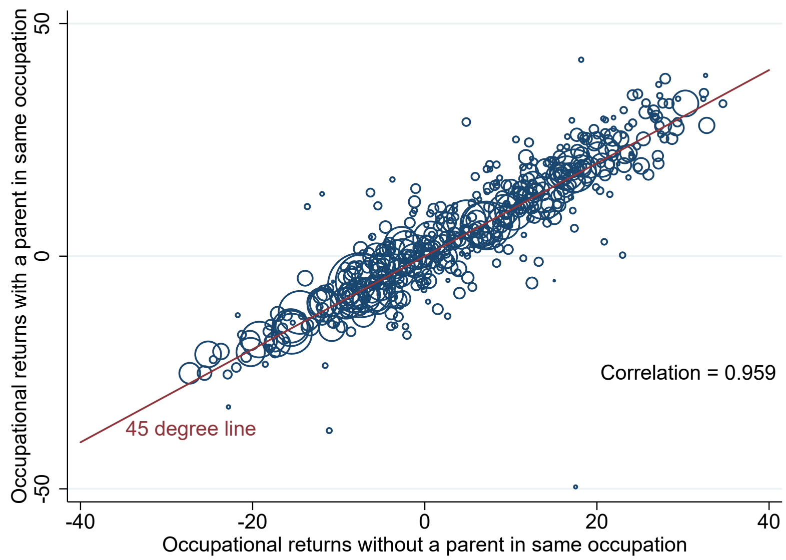
Figure A6: Intergenerational Income Mobility and the Transmission of Occupations, by Gender



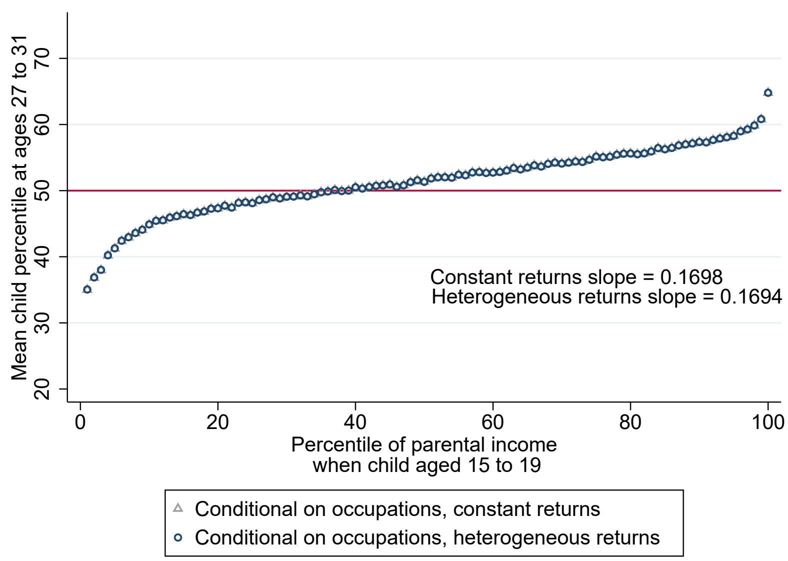
Notes: This figure replicates Figure 4, but conducts the analysis separately by child gender.

Figure A7: Heterogeneous Occupational Returns, by Parent-Child Occupation Match Status

Panel A: Occupational returns, with vs without a parent in same occupation



Panel B: Child-Parent Conditional Income Rank-rank Relationship



Notes: Panel A plots occupational returns for children with a parent in the same occupation as theirs (occupational followers), against the occupational returns for non-followers (those whose parents work in unrelated occupations). The dashed green line is the 45 degree line. Panel B plots conditional income rank-rank relationships, conditional on occupation (grey triangles), and conditional on occupation-by-follower status dummies (blue circles).

Appendix A: Data Appendix

Sample Selection The analyses are based on a linkage between the Intergenerational Income Database (IID) and six waves of the Canadian Census. We first describe the IID and then how this database was linked with Census data.

The IID is a database constructed by Statistics Canada that links the administrative tax records of children and their parents. This database includes children who were between the age of 16 to 19 in fiscal years 1982, 1984, 1991, 1996 and 2001. The IID therefore covers children born over a 23 year period, from 1963 to 1985 inclusively, but kids born in 1971, 1976 and 1981 are not included. This exclusion was caused by the sampling design of the second wave of the IID (covering birth cohorts post-1970) which had limited funds to produce child-parent links for additional cohorts. Table A1 summarizes the IID cohort structure. It also include population counts and compares them with Census count.

Table A1: The Intergenerational Income Database cohort structure

Match Year	Birth cohorts	Census	IID Count	Ratio	IID Weighted	Ratio
1982	1963 to 1966	1,723,720	1,183,614	0.687	1,517,127	0.880
1984	1965 to 1968	1,563,105	1,124,849	0.720	1,517,126	0.971
1986	1967 to 1970	1,520,745	1,155,248	0.760	1,517,127	0.998
1991	1972 to 1975	1,495,750	1,102,855	0.737	1,484,566	0.993
1996	1977 to 1980	1,570,605	1,166,879	0.743	1,558,393	0.992
2001	1982 to 1985	1,642,535	1,350,222	0.822	1,634,646	0.995

In total, there are over 7 million child-parent pairs in the IID, 5.99 million of which are unique (some birth cohorts overlap across match years, leading to duplicates entries). In the current paper, duplicate entries are excluded. In the IID, parents and children are matched together through the T1 Family File (T1FF) when children are aged 16 to 19 years old. The tax records come from the Canada Revenue Agency and includes all information contained on the tax file. The linkage process is described in details in Corak and Heisz (1999). The match rate, as measured by the IID count divided by the population count of the relevant population in the Censuses, is fairly high, ranging from 68 to 82 percent. Once weights are accounted for, the match rate of the IID count to Census count varies between 88 to 99.5 percent.

The IID records have also been linked with data from the long-form Census using a probabilistic linkage process. The long-form Census covers 1 in 5 Canadian households. It contains information about the demographic, social and economic situation of Canadians. The link has been attempted for both parents and children with Censuses 1991, 1996, 2001, 2006, 2011 and 2016. Each parent or child can be matched with more than one Census if they completed the long-form in multiple years. Depending on the birth cohort, the match rate ranges from 45-58% for the parents (average 53%), with more recent birth cohorts having a lower match rate. For the children, the match rate ranges from 42-62% (average 54%), with more recent birth cohort having a better match rate. The overall match rate for children conditional on also matching at least one of their parents is 31%. This constitutes our main analytical sample.

Children from older cohorts (those born in the early 1960s) have a higher probability of being matched in more than one Census. Using only periods were kids are observed at more or less the same age in the Census would have been too restrictive to allow for a rich documentation of occupational transmission patterns across over 500 occupations. Table A2 highlights in grey the Census waves that were used to retrieve the kids' occupation for each birth cohorts. The average age of children at the time they filled out the census is 32.5 years.

Table A2: Kids occupation and census year by birth cohorts

	Census Year					
	1991	1996	2001	2006	2011	2016
Birth years: 1963 to 1966						
Birth years: 1967 to 1970						
Birth years: 1972 to 1975						
Birth years: 1977 to 1980						
Birth years: 1982 to 1985						

For the parents, we attempted the match for each of the six Censuses and used the most recent Census in which the parent reported an occupation in order to get the most "senior" occupation. We were able to match 53% of the parents to at least one Census. On average, mothers are 57 years old when we link them to Census microdata, and fathers are 59.5 years old. Figures A1 and A2 show that the matched IID-Census subsamples produce patterns of intergenerational mobility that are very similar to those obtained from the entire IID population.

Variables Definitions Occupational codes were harmonized across years using Statistics Canada's concordance tables[17]. Occupation is identified through question 38 in the 2016 questionnaire: "What was this person's work or occupation?". It is then classified using the National Occupational Classification (NOC) or the Standard Occupational Classification (SOC) depending on the Census year. The original data for Censuses 1991 to 2006 is based on the NOC 1991, and was then converted to NOC 2011 codes. Data for Censuses 2011 and 2016 is originally based on NOC 2011 codes. We have a total of 500 harmonized occupational codes.

The main field of study variable comes from question 27 in the 2016 questionnaire. This variable was created by assigning a field of study code from the Classification of Instructional Programs (CIP). This question is only answered by respondents who obtained some post-secondary certificate, diploma or degree. For those who did not obtained a post-secondary certificate, diploma or degree, we assigned a "no diploma" category. Over the years, the categories have been slightly modified. We have harmonized them and provide the correspondence table on our website. The CIP classification can be divided into 2-digit, 4 digit and 6-digit codes. We use the 4-digit codes, which results into 446 mutually exclusive categories, including the "no diploma" category.

[17] https://www.statcan.gc.ca/en/concepts/concordances-classifications