Occupational Choice and Social Mobility

Samuel M. Solomon (Yale University)*

Job Market Paper

October 31, 2024

Please click here for the latest version.

Abstract

This paper uses population-wide administrative data from Finland to examine the underlying drivers and aggregate implications of occupational following—when children enter into their parent's occupation—for the labor market, social mobility, and total output. I first document that occupational following is a widespread feature of the labor market. Second, I show that pre-labor market multidimensional skills and educational choice statistically explain 19% of occupational following among all 43 two-digit occupations, and 53% among white collar ones. Third, I use an instrumental variables strategy to show that, conditional on pre-labor market skills and education, occupational following leads to income gains of 5.5% and fewer job separations. I then combine these mechanisms into a model of educational and occupational choice. While I find that intergenerational links play a sustained role throughout the pre-labor market and labor market years, the intergenerational transmission of occupation-specific skills is most important in driving occupational following and is productive in nature. Lastly, differences in pre-labor market multidimensional skills and educational choice—two potential policy levers—are also a key driver of occupational following and are responsible for 87% and 42% of the class gaps in white collar occupational attainment and elite occupational attainment, respectively.

*Yale University Economics Ph.D. candidate (job market paper). I am grateful to Joe Altonji, John Eric Humphries, Costas Meghir, Seth Zimmerman, and Ciprian Domnisoru for their constant guidance and support. I also thank Jason Abaluck, Jingyi Cui, Zoe Dobuler, Kristiina Huttunen, Jerry Montonen, Cormac O'Dea, Rohini Pande, and Matti Sarvimäki for their advice, as well as participants at the Yale Labor/Public Economics Prospectus Workshop, the Helsinki GSE Labor & Public Economics Ph.D. Seminar, and the Stanford SITE conference for their feedback. I thank Janne Laitila and Eric Solomon for their excellent research assistance. I thank Pengpeng Xiao for helping with data access. Finally, I thank the Cowles Foundation, The Whitney and Betty MacMillan Center for International and Area Studies, and the Institution for Social and Policy Studies (all at Yale) for generous funding. Contact email: samuel.solomon@yale.edu.

1 Introduction

Children are disproportionately likely to enter into their parent's occupation. As shown in Figure 1, nearly one in six children start their career in the same occupation as one of their parents—what I refer to as occupational following.[1] This share is around three times greater than what would be expected under random sorting into occupations.[2] While most related literature has focused on elite white collar occupations (Laband and Lentz (1992); Ventura (2023)), occupational following is a feature of all occupations across the income distribution, and is actually more common in blue collar occupations than in white collar ones.

Occupational following may have implications for intergenerational occupational mobility and occupational attainment across broad swaths of the population. Concerns regarding social mobility and access to society's highest-paying and elite jobs have recently garnered renewed interest among both academics (Chetty et al. (2023); Stansbury and Rodriguez (2024); Zimmerman (2019)) and policymakers (BMA (2023); Jiménez (2024); YLE (2021)). Such concerns have revolved around issues of equity, supply shortages, and ensuring that the best talent can enter "top" occupations. Occupational following may also reduce the level of intergenerational income mobility in society, as an occupation is typically an individual's primary source of income (Black et al. (2023)). Lastly, occupational following might also impact the efficient allocation of workers to jobs.[3]

Despite the fact that occupational following is so common, and that it has the potential to inhibit both occupational and income mobility, we have limited microeconomic evidence on the underlying reasons *why* children pursue—or do not pursue—their parent's occupation. Understanding the root causes of occupational following and occupational choice more broadly, however, is a necessary input into any policy intervention designed to increase social mobility. In parallel, it is important to understand whether such policies would come at a gain or cost to total output, which further depends on the underlying mechanisms.

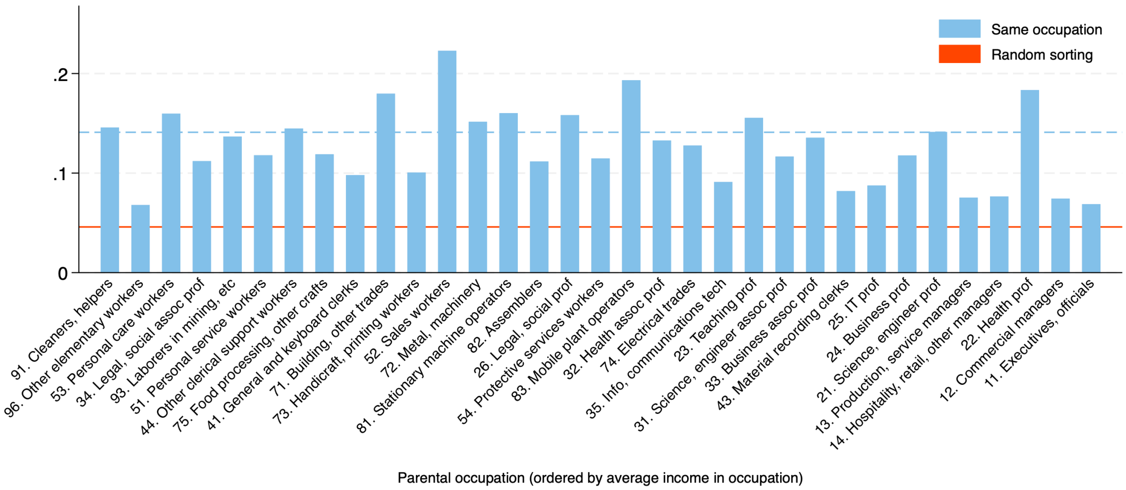
This paper studies why some children enter into their parent's occupation and others do not, and the implications of occupational persistence for the labor market, social mobility, and total output. Using rich, population-wide administrative data from Finland, this paper proceeds in two parts. First, I employ a mix of descriptive and causal inference methods to illustrate the key economic mechanisms that drive occupational following. In particular, I document the roles of pre-labor market multidimensional skills and educational choice

[1] I will use occupational following and occupational persistence interchangeably throughout the text.

[2] While these motivating statistics refer to Finland, which is the context of this paper, they are similar in the United States and other settings, as I discuss in Section 2.1.

[3] Occupational following may also affect welfare insofar as occupations have non-wage amenities and preferences over those amenities have an intergenerational component (Boar and Lashkari (2021)).

Figure 1: Rates of occupational following upon labor market entry



Notes: Figure shows the rate of occupational following by parental occupation, at the two-digit level. Occupations are ordered by mean income in occupation. Shown for all children, with respect to both parents. Blue dashed line shows the weighted mean across all occupations. Orange solid line shows random sorting benchmark (as defined in Section H.4).

and then use an instrumental variables strategy to identify the causal role of occupation-specific skills and search frictions. Second, I develop and estimate a model of educational and occupational choice, in which I embed the notion of occupational following. I use the model to determine the quantitative importance of each mechanism jointly in explaining occupational persistence and consider their implications for model-implied total output. I then use the model to decompose the sources of related class gaps in social mobility and consider the efficacy of different classes of policies in increasing mobility and whether their impacts would affect total output.[4]

In the first part of the paper, I establish five empirical facts regarding the causes and consequences of occupational following. I begin by documenting that occupational following is a widespread phenomenon as fact one. By estimating discrete choice models of occupational choice, I show that children are disproportionately likely to choose an occupation if it is their parent's most-experienced occupation. This occupational persistence holds across the entire income distribution, is relatively stronger at finer levels of occupational classification, and is stronger in blue collar occupations than in white collar ones. I then show that intergenerational occupational persistence explains a significant share of the population-wide level of

[4]Such policies include those that would equalize intergenerational skill transmission, reduce the costs of college and different college majors, and equalize labor market search frictions. Examples of these policies studied elsewhere include investments in early childhood development for skill transmission (Duncan et al. (2023)); information provisions for educational choice (Hastings et al. (2016)); and equal-hiring policies for search frictions (Bertrand et al. (2019)).

intergenerational income persistence, further warranting the investigation into occupational choice and its intergenerational component that composes the remainder of this paper.

I show that multidimensional skills at age 18 and ensuing educational choices play significant roles in explaining occupational persistence as facts two and three. I use rich data from the Finnish Defence Forces on the cognitive and socioemotional skills of nearly all males at age 18, combined with data on the field of study in their highest educational degree. As the Defence Forces data is limited to males, I restrict my analysis in this section to occupational persistence between fathers and sons. I condense the various skill measures into four skills (quantitative, verbal, social, and manual skills) for tractability and show that sons have vastly different levels of these skills depending on their father's occupation. These differences in pre-labor market multidimensional skills statistically explain a substantial share of why sons are more likely to pursue their fathers' occupation. For example, sons are 520% more likely to become a health professional (which primarily includes medical doctors) if their father is a health professional. The sons of health professionals also have quantitative skills that are on average 0.8 standard deviations greater than the average male at age 18. Conditional on all age-18 multidimensional skills, sons are 467% more likely to become a health professional, or 10% less likely relative to the baseline estimate of 520%. Furthermore, I show that conditional on their skills, sons are still disproportionately likely to pursue the education that leads to their father's occupation, further explaining why sons are disproportionately likely to enter into their father's occupation. Conditional on age-18 multidimensional skills, college attainment, and college major, sons are 116% more likely to become a health professional if their father is a health professional. Overall, differences in multidimensional skills and educational choice statistically explain nearly four fifths (78%) of why the sons of health professionals are disproportionately likely to become health professionals.

Among all 43 two-digit occupations, I find that multidimensional skills and educational choice can explain nearly one fifth (19%) of occupational persistence. Importantly, this conceals significant heterogeneity: these two pre-labor market mediators can explain a majority (53%) of white collar persistence, but a relatively small amount (9%) of blue collar persistence. Of course, the exact shares further vary across the occupational distribution and across occupational classifications, which I will characterize in great detail. Nonetheless, the main takeaway from facts two and three is that age-18 multidimensional skills and educational choice explain a significant share of occupational persistence among men, especially for white collar occupations. At the same time, a significant share remains, potentially due to differences in labor market experiences and other factors, which I elucidate next.

While facts two and three are largely descriptive, I next use an instrumental variables strategy to provide causal evidence of the underlying drivers of occupational persistence, as well as its labor market consequences. Facts four and five illustrate the role of differences in labor market returns and labor market search frictions in explaining occupational persistence. In particular, I use variation in the labor demand of the occupations of individuals' parents—conditional on occupational and local labor market fixed effects—as an instrument to estimate the causal effect of occupational following on income and job transitions. As fact four, I find that occupational following leads to income gains of 5.5% upon labor market entry relative to if the child had chosen another occupation (at the two-digit level). This effect is not driven by differences in firm access and is not statistically significantly different between sons and daughters. Lastly, in fact five, I show that occupational following reduces the likelihood of job separation upon labor market entry.

These results build on Ventura (2023), which to my knowledge is the only other paper that estimates the causal effect of occupational following on income, although it does so only for a single occupation.[5] While Ventura (2023) finds that there are positive returns to occupational following amongst the children of medical doctors, I show that there are positive returns to occupational following across all occupations on average. My findings imply that children have occupation-specific skills, conditional on the rich set of pre-labor market multidimensional skills and educational choice described earlier, that are valuable in their parent's occupation and their parent's occupation only. Unlike the multidimensional skills, these occupation-specific skills are not directly observed in the data and are valuable in single occupations only rather than across multiple occupations. Two potential examples include bedside manner for medical professions and farm-specific knowledge for agricultural work.

In the second part of the paper, I develop and estimate a structural model of educational and occupational choice to consider the mechanisms I document in the first half of the paper jointly; illuminate their interactions; assess their relative quantitative importance in aggregate; and quantify their implications for total output. Next, I use the model for another purpose. The empirical facts documented above suggest that occupational persistence can be "good," as it leads to income gains upon labor market entry of 5.5%, holding all pre-labor market skills and education constant. I use the model to explore whether—with alternative initial pre-labor market skills and education—individuals might have been better off in the long-run. I thus consider whether other policy levers still exist that would improve upward mobility at the bottom of the occupational distribution and lead to overall gains in occupational attainment and income.

[5]A series of papers in the 1980s and 1990s studied the role of occupation-specific skills in certain occupations: agriculture (Laband and Lentz (1983)), politics (Laband and Lentz (1985)), and entrepreneurship (Lentz and Laband (1990)). A common theme in each of these papers is that parents transmit occupation-specific skills to their children.

The model consists of two stages: (1) an educational stage and (2) a labor market stage. In the educational stage, children take their age-18 multidimensional skills (which depend on their parent's occupation) as given, choose whether to attend college, and if attending college, choose which major to pursue. Children do so anticipating their individual-specific value of that college major in the labor market, net of its individual-specific cost. In the second stage of the model, children enter the labor market, which I model with a search and matching framework (largely based on the model in Lise and Postel-Vinay (2020)) where they search for a job given their skills and college major. Crucially, and as a novel feature of the model, the occupational sampling distribution that individuals face in the labor market stage is conditional on their college major choice in the educational stage. Moreover, worker productivity and search frictions depend on whether one's parent is in the given occupation, consistent with the empirical evidence from above. In addition to the four mechanisms documented in the first part of the paper, I also allow children to have an idiosyncratic preference for their parent's occupation, which serves as the residual channel in the model. Finally, I estimate the model using only males, due to the aforementioned data restriction that the Defence Forces skill data is limited to males.

I first conduct a model-based decomposition of occupational persistence to illustrate the quantitative importance of each mechanism in aggregate. I find that occupation-specific skills are quantitatively the most important driver of occupational following, as they explain 44.4% of the baseline level of persistence. I also find that these occupation-specific skills represent a productive intergenerational transfer of skills: they are responsible for 0.3% of model-implied total output. Next, I find a limited role for differential search frictions and idiosyncratic preferences in driving occupational following. On the other hand, I find that differences in pre-labor market multidimensional skill endowments and differential costs of education—two actionable policy levers—explain 13% of overall occupational persistence. This share (13%) is similar to the share (19%) that I estimate in the first half of the paper—with a different method—for the role of pre-labor market skills and education in explaining occupational persistence. I also find that these two pre-labor market differences have very small differential impacts on model-implied total output (initial differences in these channels suppress total output by 0.1% combined).

I then use the model to explore whether policy levers exist to further increase social mobility, in light of the fact that children have valuable occupation-specific skills in their parent's occupation. I do this by estimating the degrees to which each mechanism explains long-run class gaps in occupational attainment and income—defined as the difference in outcomes between the children of white collar parents and the children of blue collar parents. I find that differences in pre-labor market intergenerational skill transmission are responsible for 88% and 16% of the class gap in white collar occupational attainment and elite occupational attainment, respectively. I also find that while differences in the psychic cost of college explain very little of either class gap, differences in the psychic costs of different college majors are responsible for 33% of the class gap in elite occupational attainment. This underscores the importance of considering the specific kind of education one receives, rather than just the amount. Finally, again, these pre-market differences have no significant differential consequences for overall output. These results imply that policies that promote pre-labor market skill development and reduce barriers to high-return college majors would be effective in promoting long-run social mobility, and would do so at no cost to overall output.

Relation to the literature I contribute to three main literatures: (1) intergenerational mobility, (2) the specificity of human capital, and (3) models of educational and occupational choice and search and matching models. I first contribute to the large body of research on intergenerational mobility by explicitly showing the systematic and population-wide importance of occupational choice in social mobility. Papers that have studied the role of occupational choice in intergenerational mobility have largely focused on a single occupation (Laband and Lentz (1992); Lentz and Laband (1989)).[6] Almost all of these papers study white collar occupations (and moreover, elite ones); I contribute to this literature by also studying blue collar occupations, which, as I show, are characterized by more occupational persistence than are white collar ones. Furthermore, blue collar workers make up a majority of the population (58% in my context), underscoring the importance of including blue collar workers in order to accurately assess the population-wide implications of occupational choice for intergenerational mobility.

Most of the papers that have considered all occupations have used small-scale survey data (Laband and Lentz (1983); Lo Bello and Morchio (2022)) or have been limited in their ability to speak to underlying mechanisms.[7,8] My paper is the first to consider all occupations jointly with detailed, population-wide administrative data to uncover the underlying

[6]Other papers that have focused on specific occupations include papers on doctors (Ventura (2023)), lawyers (Raitano and Vona (2021)), entrepreneurs and self-employed workers (Humphries (2022), Dunn and Holtz-Eakin (2000), Lindquist et al. (2015)), politicians (George (2019)), pharmacists (Mocetti (2016)), professionals (Aina and Nicoletti (2018); Mocetti et al. (2022)), and farmers (Emran and Shilpi (2011)).

[7]Researchers in sociology have long considered the importance of studying occupation and its relation to status (e.g., Blau and Duncan (1967)). Other papers in economics have used occupation data when income data was unavailable in broader, historical studies of economic mobility (e.g., Long and Ferrie (2013)).

[8]Montonen and Solomon (2024) use the same data as this paper, but study the role of childhood peers in occupational choice. They find that while peers play an important, causal role in individuals' occupational choices, the economic significance of peers is smaller than that of parents. However, they find that peers have a substantially greater impact when their parents work in a different occupation than one's own parents.

mechanisms of occupational following and long-run class gaps in occupational attainment more broadly. This allows me to characterize more comprehensively how occupations persist across generations throughout the entire income distribution and how that persistence varies by occupation and by the occupational classification. More importantly, I am able to more precisely identify the microeconomic determinants of occupational persistence, shedding light on the fundamental drivers of occupational choice and social mobility.

Finally, a recent set of papers has studied the degree to which college majors (Altmejd (2022); Ventura (2023)) and firms (Corak and Piraino (2011); Staiger (2023)) persist across generations. I complement and build on this work in two main ways. First, studying the college margin focuses on the distribution of workers that enter white collar jobs, implicitly excluding workers that enter blue collar jobs. I consider the entire occupational distribution, and note important heterogeneity between occupations. Second, I show that occupational persistence is stronger than both persistence in college majors and in firms. For example, Ventura (2023) studies persistence in college majors, and documents that children are on average twice as likely to choose a major when it is shared with either parent. I show that children are between 2.5 and 6.2 times more likely to choose an occupation when it is shared with either parent at a similar level of classification.[9] In terms of persistence in firms, I show that, for instance, 15% of children start their career in one of their parent's occupations, while only 6% start their career in one of their parent's firms.[10] These statistics emphasize the importance of considering occupational choice, rather than just educational choice or firm choice, when considering how economic circumstances persist across generations.

My paper also relates to the literature on the specificity of human capital, and which kinds of human capital are most rewarded in the labor market. Building from a large literature, Kambourov and Manovskii (2009) show that occupation-specific human capital is rewarded more in the labor market than is firm-specific or industry-specific human capital. A related literature has measured the returns to different types of education, finding that differences in the returns to different college majors can be as large as the college premium overall (Kirkeboen et al. (2016)). Research on the specificity of human capital also relates to research on skills, which has found that skills are multidimensional in nature (Lise and Postel-Vinay (2020)). I contribute to these literatures by showing that specific human capital (e.g., specific skills and college majors)—rather than just general human capital—persists across generations. For instance, I show that multiple dimensions of skills persist across generations, and explain more occupational persistence than does a single dimension of skill.

[9]These are at the two- and three-digit occupational levels, respectively. This corresponds with a similar level of granularity as the college major that Ventura (2023) uses; there are about 40 two-digit occupations and 120 three-digit occupations in my data and 82 majors in the classification used by Ventura (2023).

[10]Staiger (2023) finds that 5% of children in the United States start their career at their parent's firm.

This is important because it suggests that parental investment in their children's human capital is very rich in nature and that no single parental investment can capture this richness.

Lastly, my paper contributes to research on search and matching models and models of educational and occupational choice. While there is a large literature on search and matching models, Lise and Postel-Vinay (2020) made a major contribution in developing and estimating a search and matching model that allows for multidimensional skills of workers and skill requirements of jobs.[11] Few papers (e.g., Flinn and Mullins (2015)), however, have developed estimable search and matching models that include an endogenous schooling decision. Related models of educational and occupational choice include the classic model from Keane and Wolpin (1997) as well as more recent models from Altonji et al. (2012) and Todd and Zhang (2020). My paper is the first, to my knowledge, to incorporate a richly-specified college major choice that is chosen endogenously by individuals into a model with both rich multidimensional skills and rich occupational choice. By explicitly considering educational choice, and the link between skills, education, and occupation, I more accurately capture the nature of the labor market and, for instance, the degree to which pre-market differences can shape occupational choices and labor market outcomes. This is especially important for occupations such as the health and legal professions, in which pre-market differences in educational choice can entirely limit access to the occupation.

Paper structure The rest of this paper is organized as follows. In Section 2, I discuss the context and the data. In Section 3, I lay out five empirical facts related to occupational following, its causes, and its consequences. I then develop a model of educational and occupational choice, estimate the model, and discuss its implications in Sections 4, 5, and 6, respectively. I conclude in Section 7.

2 Context and data

2.1 Context

Finland's higher education system is characterized by limited monetary costs, as universities are almost entirely public and free. When students apply for a tertiary degree, they apply to a specific degree program at a specific university, as is common in many countries. Admission to degree programs largely depends on high school exams or university entrance exams. 43% of workers in Finland have a tertiary degree, while 50% of workers in the United States

[11] See also Lindenlaub and Postel-Vinay (2023) and Lindenlaub (2017) for theoretical analyses and Bandiera et al. (2024) for a recent empirical application of multidimensional matching, respectively.

do (OECD (2024)). The labor market in Finland is characterized by a similar level of employment as in the United States and each country has a similar share of their workforce in white collar occupations.[12]

While intergenerational income mobility is greater in Finland than in the United States, intergenerational occupational mobility is similar in both countries. In terms of income persistence, I estimate an overall rank-rank regression coefficient of 0.16 in my sample (Appendix Figure A1), which is lower than the corresponding estimate of 0.29 in the United States (Chetty et al. (2014)). Despite being lower than in the United States, 0.16 is still significant, suggesting that parents play an important role in the long-run economic outcomes of their children. In Appendix Figure A2, I show that while intergenerational income mobility is greater in Finland than in the United States, intergenerational occupational mobility is similar in both countries (rank-rank coefficient of 0.25 in Finland and 0.29 in the U.S.).[13]

Furthermore, Finland and the United States exhibit a similar amount of occupational following. As shown in Appendix Figure A3, a similar proportion of sons enter into their father's occupation in both countries: 11.4% in Finland and 13.6% in the United States.[14] For some occupations, there is more persistence in Finland, while for some there is more persistence in the U.S. Nonetheless, occupational persistence is not unique to Finland. Taken together, these statistics suggest that the findings in this paper are relevant for contexts outside Finland as well, such as the United States.

2.2 Data description

I use population-wide administrative data from Finland, which starts at an annual frequency in 1988 and continues until 2018. The data is especially suited for this paper because it is population-wide; it is panel in nature and linkable across time; and it contains very rich information on individuals. It contains standard demographic variables such as age, sex, language, residence, and country of birth, as well as the parents of each child, allowing me to make the relevant intergenerational link.

[12]While the occupational classifications slightly differ, both countries have just below half their workforce in white collar occupations. Finland has 46% while the U.S. has 44% (Statistics Finland (2024); U.S. Bureau of Labor Statistics (2024)).

[13]This is consistent with findings from Bjorklund and Jantti (2000), who show that the intergenerational correlation in earnings is much greater in the United States than in Finland, but that the intergenerational correlation in occupational status is similar. More recently and in a similar vein, Heckman and Landersø (2022); Landersø and Heckman (2017) show that Denmark and the United States have a comparable level of intergenerational educational mobility, despite having large differences in intergenerational income mobility.

[14]I only consider males due to the data limitations of the Panel Study for Income Dynamics, from which I draw the U.S. data for this exercise and which did not ask all women for their occupation in the early survey rounds.

I use rich data from the Finnish Defence Forces on the multidimensional skills of individuals. In contrast to the other data, this data is limited to males. Nearly all male children are required to serve in the Defence Forces at age 18, and take a test before conscription to assist in determining their placement. The test yields 11 different skill measures, composed of three cognitive measures (arithmetic, verbal, and visuospatial) and eight socioemotional measures (achievement striving, activity-energy, deliberation, dutifulness, leadership motivation, masculinity, self-confidence, and sociability). I have the cognitive measures for children born between 1964 and 1997 and the socioemotional measures for children born between 1964 and 1982.[15] To extract as much information as possible from the skill measures and reduce it to a small number of skills for computational tractability, I follow Lise and Postel-Vinay (2020) and others in using principal component analysis and a set of exclusion restrictions. This procedure yields four main measures of skills that I use in the analysis: quantitative, verbal, social, and manual. See Jokela et al. (2017) and Appendix Section C.1 for more details on these measures and on how I conduct the principal component analysis.

I complement the skill data with education data on fields of study, in addition to the standard variable indicating the simple level of educational attainment. Lastly, the labor market data includes annual work history, including the employment status, income, firm, and industry of each worker. Crucially, it also includes the occupation of each worker each year.[16] The occupation data I use includes occupations at the one-, two-, three-, and four-digit levels (as given by the International Classification of Occupations (ISCO); International Labour Organization (2023)), which are classifications in increasing order of granularity. I discuss the data and variable definitions more in Appendix C.

2.3 Sample

I use the cohorts of children born between 1975 and 1990. The Defence Force data includes the three cognitive measures for all of these cohorts, while it only includes the socioemotional measures for the cohorts born before 1982, as discussed in the previous section. I restrict the sample by dropping individuals actually in service in the Defence Forces, as they are not

[15] I have the cognitive measures for more cohorts than I have the socioemotional measures for due to a change in how the personality test was administered.

[16] While survey data typically includes occupation, the sample sizes of (even large-scale) surveys are too small to allow for a comprehensive analysis of occupations at a fine level of granularity, or are conducted only infrequently as part of a census. For instance, the United States Census collects occupation data, but only every ten years. While the American Community Survey is conducted every year, it only started in 2005. On the other hand, many governments have little or no reason to collect occupation information and so most administrative data simply does not include it. Administrative data in the U.S. that is commonly used to study labor markets, such as the Longitudinal Employer-Household Dynamics (LEHD), does not collect data on occupations.

included in the Defence Forces data (less than 1% of the overall population).

As shown in Table 1, there are over one million individuals in total in my sample: there are 16 cohorts overall, with an average cohort size of around 70,000 individuals. In terms of education, around half have earned a college degree. On average, individuals earn 21,000 Euros (€) upon labor market entry and €32,000 at age 35. I define labor market entry as the first time a worker earns €15,000 and their main activity for the year is employment.[17] Note that 13% of individuals are in the same two-digit occupation as either of their parents upon labor market entry, while 13% are at age 35 as well. By age 35, however, 33% of individuals will have at some point in their career been in the same occupation as either of their parents.[18] In all cases, occupational following is significantly more common than is firm following. Finally, the Defence Forces sample exhibits a similar amount of occupational following as the overall sample.

2.4 Discussion of skill measures

Table 2 shows the average skills of individuals in a selection of diverse occupations across the income distribution. The skill measures capture heterogeneity across broad, one-digit occupations, as well as heterogeneity within one-digit occupations across finer, two-digit occupations. For instance, professionals (code 2) tend to have high levels of quantitative and verbal skills, compared to both other white collar occupations (e.g., managers, code 1) as well as blue collar occupations (codes 4-9). Within professional occupations, there still remains differences across finer two-digit codes. For instance, health professionals—compared to teaching professionals—tend to have higher levels of quantitative and verbal skills, but lower levels of social skills. Managers—compared to professionals—tend to be characterized by greater levels of social skills, despite having lower levels of quantitative and verbal skills.

On the other hand, blue collar occupations on average are characterized by greater levels of manual skills. As among the white collar occupations, rich heterogeneity exists within broad occupational classes. For example, while both protective services workers (e.g., firefighters and police officers) and stationary machine operators have a similar level of manual skills, protective services workers tend to have a much greater level of social skills. Moreover, differences exist within one-digit occupations for blue collar occupations too: for example, protective services workers tend to have greater levels of all skills compared to personal services workers (e.g., bartenders and hairdressers). Finally, Appendix Table A2 reports the

[17] See Appendix C for more detail on the main activity variable. This definition is similar to the definition of "first stable job" in other papers, such as Kramarz and Skans (2014); San (2021); Staiger (2023).

[18] See Appendix Table A1 for these rates at different levels of occupational classification.

Table 1: Sample summary statistics

	All	Defence Forces
Demographics		
Year of birth	1982	1977
Age at labor market entry	25	25
Male	0.51	1.00
Urban	0.77	0.72
Native born	0.87	0.98
Finnish language	0.84	0.96
Adopted	0.01	0.01
Can link with parents	0.93	1.00
Parent characteristics		
Father income	24,103	22,294
Mother income	17,248	16,013
Education outcomes		
College degree	0.498	0.444
Labor market outcomes: labor market entry		
Income (€)	21,204	20,540
Same occupation as either	0.132	0.121
Same occupation as father	0.075	0.097
Same occupation as mother	0.077	0.049
Same firm as either	0.052	0.068
Same firm as father	0.034	0.052
Same firm as mother	0.023	0.021
Labor market outcomes: age 35		
Income (€)	32,045	40,080
Same occupation as either	0.133	0.124
Same occupation as father	0.079	0.104
Same occupation as mother	0.075	0.043
Ever same occupation as either	0.331	0.335
Ever same occupation as father	0.206	0.281
Ever same occupation as mother	0.200	0.149
Same firm as either	0.037	0.052
Same firm as father	0.025	0.043
Same firm as mother	0.017	0.018
Ever same firm as either	0.145	0.167
Ever same firm as father	0.090	0.129
Ever same firm as mother	0.074	0.061
Observations	1,137,039	127,859

Notes: Sample means shown. First column shows entire population, while second column is limited to Defence Forces sample. Occupation defined at two-digit level.

Table 2: Nature of skill measures

	Skill Measures			
	Quant.	Verbal	Social	Manual
13. Production managers (0.01)	0.39	0.34	0.38	-0.08
22. Health professionals (0.02)	0.81	0.84	0.11	-0.39
23. Teaching professionals (0.04)	0.66	0.66	0.25	-0.44
51. Personal services workers (0.04)	-0.18	-0.14	0.08	-0.14
54. Protective services workers (0.02)	0.08	0.20	0.31	0.12
81. Stationary machine operators (0.07)	-0.22	-0.23	-0.08	0.16

Notes: Selection of two-digit occupations shown. Sample means shown and include all children in occupation upon labor market entry. Skill measures are standardized to mean zero and standard deviation one within cohort. "Quant." is abbreviation for quantitative. Sample frequency in occupation shown in parentheses.

correlation matrix between the four skills. While the two cognitive skills are moderately correlated (correlation of 0.68), all other pairs of skills have a very low level of correlation. This further suggests that the skills capture different productive attributes.

In terms of occupational skill requirements, I follow Fredriksson et al. (2018) and use the average skills of experienced workers in the given occupation, thus assuming that workers that are experienced in a given occupation are relatively well matched for the skill requirements of that occupation.[19] Lastly, I follow recent papers on multidimensional skill mismatch (e.g., Fredriksson et al. (2018)) and consider two different measures of skill mismatch: (1) a horizontal measure which proxies for how close an individual's mix of skills aligns with the mix of skills required in their occupation and (2) a vertical measure which proxies for how close an individual's average skill level aligns with the average skill level required in their occupation. See Appendix Section C.3 for more details regarding these measures.

Lastly, the skill measures have strong predictive power in terms of the sorting of children into broad one-digit occupations (Appendix Table A3) as well as amongst more specific two-digit occupations (Appendix Table A4). The skill measures also have predictive power in terms of income (Appendix Table A5). They predict income both across occupations and workers and within occupations and workers. The fact that these skill measures vary across occupations and workers and that they have strong predictive power in terms of both occupational choice and income implies that the skill measures capture real differences in valuable human capital.

[19] I focus on workers with at least five years of experience in the given occupation. Fredriksson et al. (2018) considers workers with three years of tenure in the current job when defining the skill requirements.

3 The causes and consequences of occupational following: five empirical facts

I now document a set of descriptive and causal empirical facts that illustrate the main economic mechanisms and labor market impacts of occupational following. In addition to providing empirical evidence for the causes and consequences of occupational following, this section also serves to motivate the structural model that I introduce in the next section.

3.1 Fact I: There is strong intergenerational persistence in occupational choice, across the entire income distribution

There is strong intergenerational persistence in occupational choice for nearly all occupations across the income distribution, and this persistence occurs at both broad and fine levels of occupational classification. In order to estimate how occupations persist across generations, I estimate the following discrete choice model:

$$y_{ij} = \beta_{0j} + \beta_1 \cdot \text{parent in occupation j} + \beta_{2j} \cdot X_i + \epsilon_{ij}, \tag{1}$$

where y_{ij} is an indicator if individual i chooses occupation $j \in \mathcal{J}$ when they enter the labor market; parent in occupation j is an indicator if their parent's most-experienced occupation is occupation j; and X_i is a vector of demographic controls. The main coefficient of interest is β_1, which represents how much more likely an individual is to choose an occupation if it is their parent's most-experienced occupation. I first estimate a single β_1 for all occupations to obtain an average estimate across all occupations. In a secondary specification, I let the coefficient β_1 vary by the occupation and estimate it separately for each occupation $j \in \mathcal{J}$, which I denote as β_{1j}. Lastly, I consider all children with respect to the primary earner in the household in the main specification.

Table 3 shows that children are significantly more likely to choose an occupation if it is their parent's occupation, across all levels of classification and across both blue and white collar jobs. As shown in Column (1), children are 14.5 percentage points more likely to enter a given occupational class (white collar or blue collar) if their parent is in it. Given the control mean likelihood of 41%, this represents a 35% increase in the likelihood of following one's parent into the same occupational class. As one considers finer occupational categories, the relative likelihood of choosing the same occupation as one's parent increases. At the one-digit level, as shown in Panel B, children are 5.84 percentage points more likely to choose an occupation if it is their parent's most-experienced occupation, representing a relative increase of 58%. The remaining three panels show that children are 3.80, 2.92, and 2.66 percentage points more likely to choose a two-, three-, or four-digit occupation if it is their parent's occupation, respectively. While the percentage point estimate is decreasing, the relative size of occupations at finer levels of classification decreases faster. This results in an increase in the relative probability of choosing the same occupation as one's parent as the classification becomes finer. Children are 155%, 363%, and 1065% more likely to choose the same two-, three-, and four-digit occupation as their parent, respectively.

Heterogeneity across occupations Table 3 shows that while occupational persistence is strong for both white collar and blue collar occupations, it is significantly stronger for blue collar occupations at the two-, three-, and four-digit levels. At the two-digit level, the children of white collar workers are 114% more likely to choose their parent's occupation, while the children of blue collar workers are 188% more likely to choose their parent's occupation. The differences are even starker at the three- and four-digit levels.

To determine the persistence of each individual occupation, I estimate equation 1 allowing the amount of persistence β_1 to vary by occupation j to obtain β_{1j}. I then divide β_{1j} by the control mean to estimate the percentage increase in the likelihood of choosing occupation j if one's parent is from occupation j, and plot that estimate for each occupation in Appendix Figure A4. As shown in the figure, children are more likely to choose their parent's occupation in nearly all occupations.[20]

After initially documenting the average amount of intergenerational occupational persistence, I have now shown that occupational persistence is a feature of almost all individual occupations across the income distribution rather than just a feature of a select few. Moreover, with the exception of the one-digit level, children are roughly two times more likely to choose the same occupation as their parent in blue collar occupations compared to in white collar occupations. As discussed in Section 1, most related papers on intergenerational occupational persistence have focused on high-paying white collar occupations, such as doctors and lawyers (e.g., Laband and Lentz (1992); Lentz and Laband (1989); Raitano and Vona (2021); Ventura (2023)). However, this narrow focus misses an important part of the story: not only is following also very common in other white collar occupations and in blue collar occupations, it is *more* common in blue collar occupations. Taken together, this warrants a systematic, economy-wide investigation into the microeconomic drivers of occupational persistence as well as its aggregate consequences, as done in this paper.

[20] For reference and for the reader's curiosity, Appendix Table A6 lists all occupations in decreasing order of persistence, by occupational classification.

Table 3: Intergenerational occupational persistence

	Choose occupation j		
	All	BC	WC
	(1)	(2)	(3)
Panel A. Class level			
Parent in occupation j	0.145 (0.0019)		
Observations (millions)	0.5		
Mean	0.4105		
Percent increase	35%		
Panel B. One-digit level			
Parent in occupation j	0.0584 (0.0009)	0.0541 (0.0011)	0.0639 (0.0013)
Observations (millions)	2.3	1.5	0.8
Mean	0.1015	0.0972	0.1106
Percent increase	58%	56%	58%
Panel C. Two-digit level			
Parent in occupation j	0.0380 (0.0005)	0.0467 (0.0008)	0.0272 (0.0007)
Observations (millions)	9.8	6.0	3.8
Mean	0.0244	0.0248	0.0239
Percent increase	155%	188%	114%
Panel D. Three-digit level			
Parent in occupation j	0.0292 (0.0009)	0.0423 (0.0015)	0.0142 (0.0011)
Observations (millions)	5.9	3.1	2.7
Mean	0.0081	0.0095	0.0064
Percent increase	363%	444%	222%
Panel E. Four-digit level			
Parent in occupation j	0.0266 (0.0009)	0.0391 (0.0014)	0.0116 (0.0009)
Observations (millions)	18.1	9.2	8.9
Mean	0.0025	0.0031	0.0019
Percent increase	1065%	1280%	604%

Notes: Results from estimating equation 1 shown. Job upon labor market entry of all children considered, with respect to parent that is primary earner. Demographic controls include sex, age, year of birth, native born, and language. Percent increase equals coefficient of interest divided by control mean shown. Shown for all occupations, and then separately for blue collar occupations and white collar occupations. Class level, one-digit, and two-digit specifications limited to 1975-1979 cohorts for computational feasibility. Three-digit and four-digit specifications limited to 1978 cohort only for computational feasibility. Standard errors in parentheses.

Other heterogeneity In Appendix Table A7, I show how the intergenerational link varies with the gender of the parent and the child. While children are more likely to choose an occupation if it is held by either parent, they are significantly more likely to choose an occupation if it is held by their same-sex parent rather than by their opposite-sex parent.

Appendix Table A8 shows that occupational following is more common for those born in rural areas, for those that are native-born, and for those that are biological rather than adopted children. In Columns (2) through (4) of Appendix Table A8, I consider selection into following on the basis of skills and education, which requires that I limit the exercise to the Defence Forces sample. Overall, attending college and having more cognitive skills are associated with less occupational following, while having more manual skills is associated with more occupational following; this pattern implies that there is negative selection into occupational following on observable characteristics overall. Importantly, this negative selection is driven by occupational following in blue collar occupations; in white collar occupations, there is positive selection on the same set of observable characteristics.[21,22]

Link to intergenerational income persistence Appendix Section D shows that intergenerational occupational persistence is an important driver of intergenerational income persistence. First, I show that occupational followers experience significantly less upward income mobility. For children born to parents at the 25th percentile of the income distribution, the average income rank for children that enter their parent's occupation is 41.3, while it is 46.8 for those that do not enter their parent's occupation—5.5 ranks larger. Second, I show that occupational following, and occupational choice more broadly, can explain a substantial share of the population-wide level of income persistence. These findings suggest that occupational choice is not only quantitatively important for intergenerational occupational mobility, but it is also quantitatively important for intergenerational income mobility.

3.2 Fact II: Age 18 multidimensional skills play an initial mediating role in explaining occupational persistence

Now that I have shown that occupational persistence is a common feature across all occupations and a key driver of income persistence, I turn to discussing the microeconomic

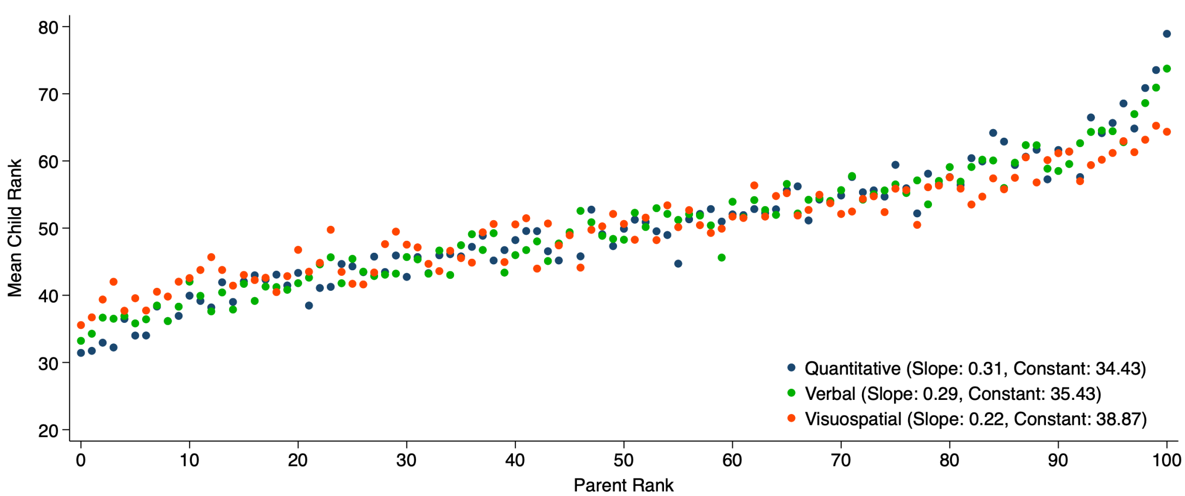
[21] For children of blue collar workers, attending college and having more cognitive skills is negatively associated with occupational following. On the other hand, the opposite is true for the children of white collar workers. This difference implies that there is negative selection into occupational following on observable characteristics for blue collar workers, while there is positive selection for white collar workers.

[22] For all occupations, I find that sons are more likely to follow their father if they have skills that align with their father's occupation: a one standard deviation decrease in horizontal skill mismatch is associated with a 1.71 percentage point, or 15.9%, increase in the likelihood of occupational following.

determinants of occupational persistence. As fact two, I show that age 18 multidimensional skills vary by a child's parental background and that these differences in skills play an initial mediating role in occupational persistence. I start by estimating the intergenerational rank-rank relationship in cognitive skills. Note that given the cohort coverage of the skill measures as described earlier, I have a sufficient number of cohorts of the cognitive skill data—but not the socioemotional skill data—to estimate this relationship directly. I therefore estimate this relationship indirectly using imputed socioemotional skills for the parents. Also note that, again as discussed earlier, I only have skill measures for males, and so this fact will only refer to the degree to which skills explain occupational persistence from fathers to sons.

Figure 2 shows the rank-rank relationship for each cognitive skill. The rank-rank slope for quantitative skills is 0.31, which is larger than the rank-rank slope for income in Finland (0.19 in this sample, as discussed above) and roughly equal to the rank-rank slope for income in the United States (Chetty et al. (2014)). For example, sons born to fathers at the 95th percentile of quantitative skills in their cohort are on average in the 70th percentile in their own cohort. The comparable estimate for the rank-rank relationship in income is the 55th percentile, highlighting the strength of the intergenerational relationship in quantitative skills. Verbal skills and visuospatial skills also persist strongly across generations. Appendix Figure A5 illustrates this relationship using imputed socioemotional skill measures and shows that while the relationship is weaker compared to the cognitive measures, it is still substantial.

Figure 2: Intergenerational multidimensional skill transmission



Notes: Father cognitive skills taken directly from Defence Force data for the sample for which this is feasible. Data is only available for males. Skills are standardized individually within cohort to mean zero and standard deviation one. Child rank defined within child cohort. Parent rank defined within child cohort.

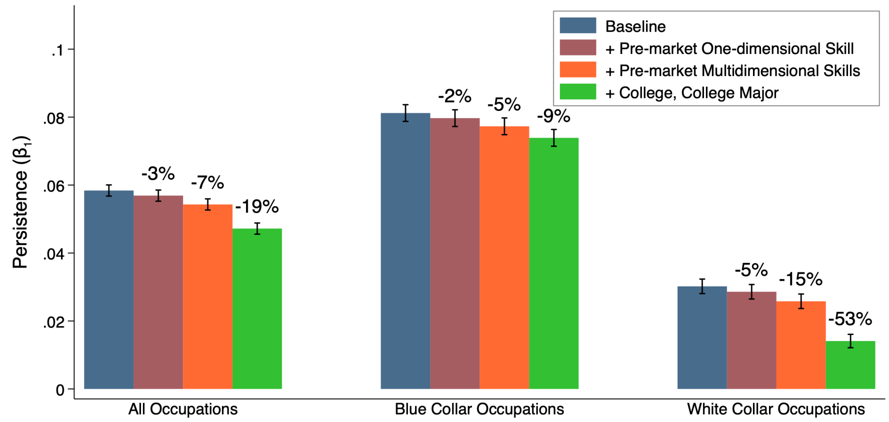
Given children's multidimensional skills vary by their parental background, I now estimate to what degree these differences explain why occupations persist across generations. To do so, I conduct a statistical decomposition of occupational persistence, building on the discrete choice framework that I introduced in Section 3.1. In particular, I now condition on the multidimensional skills of each individual by adding $\beta_{3j} \cdot \mathbf{x_i}$ to Equation 1, where $\mathbf{x_i}$ is individual i's vector of multidimensional skills. I then consider how the estimate of persistence, β_1, changes relative to the baseline case.

Focusing discussion on the two-digit level, the first bar in Figure 3 shows that sons are 5.84 percentage points (or 248%) more likely to choose an occupation if it is their father's most-experienced occupation (see Appendix Table A9 for all underlying estimates). In the second bar, I condition on a single dimension of skill—taken as the simple average of all four skills—and show that conditional on the single dimension of skill, children are 5.69 percentage points more likely to choose an occupation if it is their parent's occupation. This implies that a single dimension of skill can explain 3% of overall occupational persistence at the two-digit level. In the third bar, I condition on multiple dimensions of skill—conditioning on all four measures individually—and show that conditional on multidimensional skills, children are 5.43 percentage points more likely to choose an occupation if it is their parent's occupation. Pre-labor market multidimensional skills overall can thus explain 7% of occupational persistence at the two-digit level. This underscores the importance of having multiple dimensions of skills, as including multiple dimensions explains significantly more than does a single dimension. Considering only a single dimension of skill, therefore, would incorrectly attribute part of occupational persistence to the residual.

3.3 Fact III: Conditional on skills, children still make vastly different educational choices

Conditional on multidimensional skills, children still make very different educational choices, which further explains occupational persistence. In the fourth bar of Figure 3, I additionally control for whether children attend college and, if so, what their college major is. The college major classification consists of 101 different majors, such as history, medicine, and psychology (see Appendix C). In particular, I add $\beta_{4j} \cdot \mathbf{s_i}$ to Equation 1, in addition to the multidimensional skill controls, where $\mathbf{s_i}$ is a vector of indicators for each college major, with each element equal to one for individual i's college major and zero otherwise. I then consider how the estimate of persistence, β_1, changes relative to the baseline case. College major explains an additional 12% of occupational persistence at the two-digit level. In total, multidimensional skills at age 18 and subsequent college major choice explain nearly one fifth (19%) of occupational persistence at the two-digit level.

Figure 3: Intergenerational occupational persistence: the roles of skills and education



Notes: First (blue) bar in each group shows estimates of β_1 from equation 1. Second (red) bar includes control for one-dimensional skill, third (orange) bar includes individual controls for each multidimensional skill, and fourth (green) bar also includes control for college major (including category of no college attainment). Number above bar shows change in magnitude relative to baseline estimate. Shown for all occupations, and then separately for blue collar occupations and white collar occupations. Job upon labor market entry of sons considered with respect to fathers, due to the fact that skill data is limited to males. Demographic controls include age, year of birth, native born, and language. Shown at two-digit level. Limited to 1975-1979 cohorts for computational feasibility. 90% confidence intervals are shown.

As shown in Figure 3, the degree to which skills and educational choice explain occupational persistence varies significantly by occupational class. For white collar occupations, multidimensional skills explain 15% of the persistence and educational choice explains another 38%, for a total of 53%. This is over five times the amount of persistence that is explained by these factors for blue collar occupations, where skills and education only explain 9% of persistence in total.

There is considerably more heterogeneity than is revealed through a simple comparison of white collar and blue collar occupations alone. In Appendix Table A10, I show the decomposition at the one-digit level by occupation. While skills and education can explain a substantial share for all white collar occupations, they can explain almost three times the amount of persistence amongst professionals (66% in total) than amongst managers (23% in total). Furthermore, there is still substantial heterogeneity when looking within broad one-digit occupations, such as the professionals occupation. For example, Appendix Table A11 shows that skills and education can explain twice the amount of persistence in the health and legal professions (78% and 65%, respectively) compared to the business and IT professions (43% and 30%, respectively). The fact that the degree to which pre-labor market mechanisms, such as educational choice, can explain occupational persistence varies by occupation is consistent with and builds on the work of Altonji et al. (2012) and others, that find that the link between college major and occupation is stronger for some occupations than for others. This source of heterogeneity—the fact that some occupations more than others require specific skills and education—suggests that interventions to improve mobility into restrictive occupations must occur at a young age. Once children enter the labor market, there is relatively limited scope to affect choice of these occupations. With that being said, skills and educational choice cannot explain the entirety of the intergenerational relationship for any occupation. This suggests that other mechanisms, such as labor market returns or search frictions, differ by parental background. I turn to these mechanisms in the next two facts, in Sections 3.4 and 3.5.

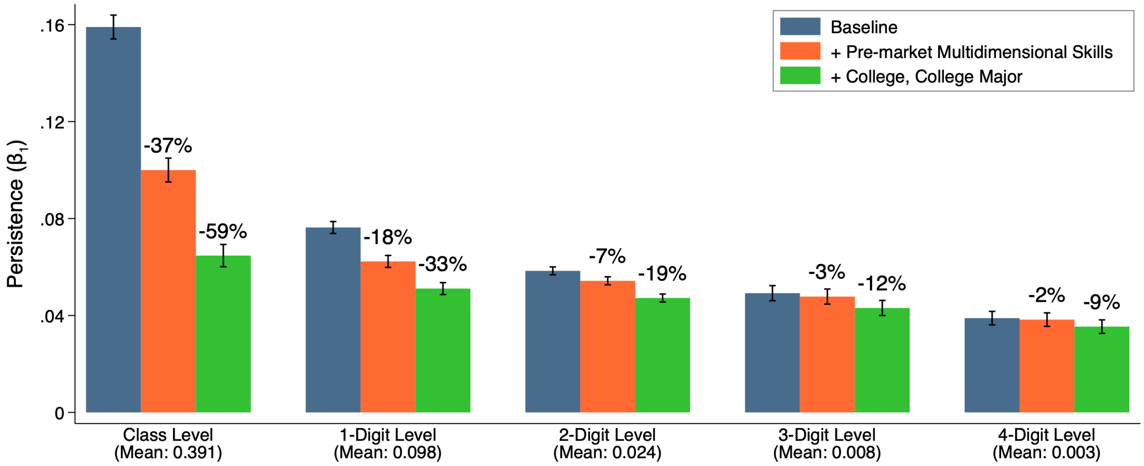
To conclude this section, I show that the explanatory power of skills and education depend on the coarseness of the occupational classification. Figure 4 shows that multidimensional skills and education can explain 59% of all occupational persistence at the class level. This percentage decreases, though, as one considers finer occupational classifications. Multidimensional skills and education explain 33%, 19%, 12%, and 9% of persistence at the one-, two-, three-, and four-digit levels, respectively. Thus, while sons are over 16 times more likely to choose a four-digit occupation if it is their father's occupation, one can barely explain such fine occupational persistence with multidimensional skills and education alone. Other channels, such as occupation-specific skills and idiosyncratic preferences, must play a larger role in persistence at these finer levels. The fact that the degree to which these mediators explain occupational persistence varies so greatly is an important finding: it suggests that multidimensional skills and the college major choice are important at broader levels of occupational classification (such as whether the child of a teacher becomes a teacher), but essentially unimportant at finer levels of occupational classification (such as whether the child of a *primary school* teacher becomes a *primary school* teacher).

Given these results, the remainder of the paper will focus largely on the two-digit occupational classification. This decision is further supported by the fact that the two-digit level explains significant wage variation overall and that most of the wage variation explained by occupational choice is captured at the two-digit level (Appendix Table A12). In addition, the share of the rank-rank relationship in income that can be explained by occupational choice is mostly captured at the two-digit level as well (Appendix Table A13).

3.4 Fact IV: Conditional on pre-market skills and education, occupational following increases income upon labor market entry

I now turn from pre-labor market mechanisms to labor market mechanisms that drive occupational following, and estimate the effect of occupational following on labor market outcomes. I begin by considering the ordinary least squares (OLS) difference in labor market outcomes.

Figure 4: The roles of skills and education by occupational classification



Notes: See notes from Figure 3. All occupations included. In contrast to Figure 3, the decomposition is done separately for each occupational classification, as indicated on horizontal axis. Control mean shown in parentheses. One-dimensional skill not included for expositional purposes. See Appendix Table A9 for all underlying estimates.

A potential drawback of the OLS estimate, however, is non-random selection into who follows their parents in the first place. On one hand, children might follow their parents because of an idiosyncratic preference for their parent's occupation, which could lead to them accepting a lower wage. On the other hand, if children have occupation-specific skills in their parent's occupation, one would expect them to earn more in their parent's occupation than if they had chosen another occupation.

To address this concern, I use idiosyncratic variation in the labor demand in the occupation of one's parent as an instrument to estimate the causal effect of occupational following on labor market outcomes. The exclusion restriction is that, conditional on parental occupation fixed effects and parental local labor market (region-year) fixed effects, the regional hiring rate in the occupation of one's parent affects the child's labor market outcomes only through whether or not they enter into their parent's occupation in the first place. This type of instrument—residual idiosyncratic variation in labor demand upon labor market entry—has been used in other contexts as well, including Staiger (2023)'s analysis of the returns to entering into one's parent's firm and Arellano-Bover (2024)'s analysis of the returns to starting one's career at a large firm.

More formally, I estimate the following two stage least squares regression equations. The first stage is given by the following:

$$S_{it} = \gamma_0 + \gamma_1 Z_{ort} + \theta_o + \lambda_{rt} + X_i + \nu_{it}, \tag{2}$$

where t is the year in which child i enters the labor market (defined as in Section 2.3); S_{it} is an indicator for whether child i chooses the same occupation as their parent; Z_{ort} is the regional hiring rate (defined as the raw count of new hires) in their parent's occupation o at time t in region r; θ_o is a fixed effect for their parent's occupation o; λ_{rt} is a fixed effect for the local labor market (region-year) of their parent's job; and X_i is a vector of child demographic controls (age, sex, native born, language, and parental income).

The second stage is given by:

$$y_{it} = \beta_0 + \beta_1 S_{it} + \theta_o + \lambda_{rt} + X_i + \epsilon_{it}, \tag{3}$$

where y_{it} are the labor market outcomes of interest for child i (income and job transitions); θ_o and λ_{rt} are again fixed effects for their parent's occupation and local labor market, respectively; and X_i are the same set of controls as in the first stage. In the main specification, I consider all children with respect to the primary earner in the household.

Main results The main results are shown in Table 4. The instrument has a strong first stage, as shown in Column (1), with a partial F-statistic of 677. A one percent increase in the hiring rate in the occupation of one's parent increases the likelihood that the child enters that occupation by 0.13%.[23] Starting in Column (3) with the main specification, I now consider the second stage estimates. Entering into one's parent's occupation upon labor market entry leads to an increase in income of 5.5%, compared to entering into another occupation. This estimate is both statistically and economically significant. Note that the IV estimate is larger than the comparable OLS estimate of 1.8%, as shown in Column (2). This is consistent with the overall negative selection of occupational followers on observable characteristics, as shown earlier in Section 3.1.[24]

Columns (4) through (6) of Table 4 show that this estimate is stable across alternative specifications. Column (4) addresses the possible concern that children might be more likely to enter into their parent's occupation if their parent's occupation is growing. The specification in Column (4) thus includes the regional growth rate in hiring in the parent's occupation as a control. In Column (5), I address the concern that children might manipulate their timing of entry into the labor market by using the hiring rate from the previous time period.

[23] This is the interpretation of the point estimate of the first stage. An increase in the hiring rate in one's parent's occupation of 10,000 raw hires leads the child to be 11.0 percentage points more likely to enter into their parent's occupation. Given a mean hiring rate of 0.097 (in 10,000's) and following rate of 0.083, this implies that a 1% increase in the hiring rate leads the child to be 0.13% more likely to enter into their parent's occupation.

[24] Similarly, Ventura (2023) also finds there to be negative selection in her analysis of the returns to medical school, on the basis of whether one's parent is a doctor. Staiger (2023) also finds negative selection into firm following.

Table 4: The effect of occupational following on income

	Same occ.	Log income						Placebo test	
	First stage	OLS	IV					First stage	Reduced form
	(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)
Hiring rate	0.110***							0.000	-0.003
	(0.005)							(0.004)	(0.003)
Same occupation		0.018***	0.055*	0.065**	0.086**	0.074**			
		(0.001)	(0.031)	(0.031)	(0.038)	(0.036)			
Partial F-statistic	677		677	682	468	498		0	
First stage magnitude	0.13%								
Following rate	0.083	0.083	0.083	0.084	0.083	0.083		0.083	
Observations	621,444	621,444	621,444	617,616	613,261	621,366		606,163	606,163
Additional controls	None	None	None	Growth	None	Coarse		None	None
Time of hiring rate	Entry	Entry	Entry	Entry	Prior	Entry		Entry	Entry

Notes: All children included with respect to parent that is primary earner. Occupation taken at labor market entry. Two-digit occupation shown. Hiring rate shown in 10,000's. Demographic controls include age, sex, native born, language, and sex and income of primary earner. "Growth" time control specified as regional growth rate in hiring in occupation. "Prior" time of hiring rate refers to hiring rate in year prior to labor market entry. "Coarse" control includes hiring rate for one-digit occupation in local labor market. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Finally, in Column (6), I address the concern that there still might be correlated shocks in labor demand across related occupations that are not absorbed by the occupation fixed effects or the local labor market fixed effects. For example, there might be shocks to certain professional occupations (one-digit code 2)—but not all professional occupations—within a local labor market that might in and of itself lead individuals to earn more. In Column (6), I thus control for the regional hiring rate in that location in the one-digit occupation from which the parent's two-digit occupation belongs (for example, I control for the hiring rate among one-digit code 2 overall if one's parent is in two-digit code 21). This specification is the strictest one, as I am isolating idiosyncratic variation in the parent's occupation net of the variation from the most related occupations. Overall, I find that the main estimate of the effect of occupational following on income is stable across all specifications. Finally, in Appendix Table A14, I show that the effect on income is sustained over at least 3 years.

Heterogeneity I find no statistically significant differences by occupational class (Appendix Table A15) or by gender (Appendix Table A16). I find that the effects are larger at finer levels of occupational classification (Appendix Table A17).

Placebo test and more robustness Finally, I consider a placebo test to assess whether the variation in labor demand I am using is truly isolated to the parent's occupation and does not reflect correlated shocks across related occupations. While the local labor market fixed effect in the main specification absorbs shocks that are common across occupations within a local labor market, it will not absorb shocks that are unique to a certain set of multiple—but not all—occupations in a local labor market. As a placebo test for this potential issue, I replace the hiring rate in the occupation of one's parent with the hiring rate from a related yet different occupation. In particular, I randomly use the hiring rate from another two-digit occupation in the same one-digit occupation. For example, consider occupation code 51 (personal service workers). I replace the hiring rate with a randomly chosen different two-digit code that has the same one-digit code: I randomly replace the hiring rate from 51 with the hiring rate from 52 (sales workers), 53 (personal care workers), or 54 (protective services workers). Each of these occupations are in the same one-digit occupation 5 (service and sales workers) and are relatively similar in nature, yet still distinct, and thus this placebo test will examine whether there are correlated shocks across similar occupations that are driving the results or whether the identifying variation is truly isolated to the specific occupation of one's parent. As shown in Columns (7) and (8) of Table 4, this placebo test yields no significant first stage or reduced form, respectively.[25] This is further evidence that the identifying variation driving the causal effect of occupational following on income is truly isolated to the *specific* occupation of one's parent.

As a final robustness check, I consider whether the results could be due to entering into one's parent's firm. In Appendix Table A19, I show that both the OLS and IV estimates do not significantly change when controlling for being in the same firm as one's parent.

Interpretation Given the placebo test and robustness checks above, I interpret the fact that children earn more when they enter their parent's occupation as evidence that they have an occupation-specific skill or advantage that is valuable in their parent's occupation and their parent's occupation only. In contrast to the pre-labor market multidimensional skills (i.e., the quantitative, verbal, social, and manual skills) described earlier, these occupation-specific skills are not directly observed in the data and are valuable in single occupations only rather than across multiple occupations. Potential examples include bedside manner for medical professions, culinary skills for food services, farm-specific knowledge for agricultural work, and carpentry skills for construction.

3.5 Fact V: Occupational following also leads to fewer job separations

Next, as fact five, I show that occupational following also leads to fewer job separations. Here, I consider whether individuals are more likely to stay in their job when they enter their parent's occupation or if they are more likely to experience a transition into unemployment or into another job. As shown in Column (2) of Table 5, I find that occupational following significantly reduces the likelihood of job separation by 9.6 percentage points. As shown in Appendix Table A20, this effect is robust across all the alternative specifications (although the estimate becomes noisier) and placebo tests considered above in Section 3.4. This implies that occupational following is also associated with reduced search frictions, in addition to the mechanisms that I have documented thus far.

3.6 Alternative mechanisms

As I will discuss in Section 6, the mechanisms included in the model—which reflect the empirical facts above—can explain 64.3% of two-digit occupational persistence. Given that occupational choice and intergenerational mobility are such complex processes, one would

[25] Appendix Table A18 shows that the placebo test results are consistent across all occupational classifications.

Table 5: The effect of occupational following on job transitions

	Job transitions		
	Stay	E2U	E2E
	(1)	(2)	(3)
Panel A. IV estimates			
Same occupation	0.023	-0.096**	-0.056
	(0.068)	(0.047)	(0.085)
Panel B. OLS estimates			
Same occupation	0.003	0.004**	-0.011***
	(0.002)	(0.002)	(0.003)
First stage F-statistic	717	394	385
Outcome mean	0.580	0.084	0.249
Observations	594,378	362,878	326,332

Notes: Main specification (Column (3) of Table 4) shown. Column (1) refers to no change in job, Column (2) refers to transition into unemployment, and Column (3) refers to transition into a different job. All children included with respect to parent that is primary earner. Occupation taken at labor market entry. Two-digit occupation shown. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

expect many mechanisms to contribute. I thus do not anticipate to be able to capture all occupational following with a limited number of economic mechanisms, in a single tractable and estimable model.[26] With that being said, other mechanisms might still impact occupational following, which I now discuss.

Firms I do not consider differential access to firms in my model. Staiger (2023) finds that 5.2% of children start their career in their parent's firm in the United States. I find that a similar share (3.7%) of children start their career in their parent's firm in Finland. However, most (83.0%) occupational followers in Finland are not firm followers. Staiger (2023) also finds that firm followers in the United States earn 19% more, and that this is almost entirely driven by access to higher-paying firms overall (i.e., firms with higher AKM fixed effects (Abowd et al. (1999))). As shown in Appendix Table A21, I find a similar result

[26]I can explain a similar share of occupational following as other related papers can explain in their economic outcome of interest. For example, Bolt et al. (2024) explore mechanisms that lead to the integenerational transmission of earnings, and find that the mechanisms they consider (education, skills, parental investments, and family background) can explain 55-68% of the overall earnings persistence. Lise and Postel-Vinay (2020) conduct a similar decomposition to mine, but instead of considering the determinants of occupational following, they consider the determinants of the overall social value of economic output. They can explain about 84% of the overall social value of output with their observable characteristics.

in Finland: using ordinary least squares, I find that firm followers enter into firms that pay all workers 5.6% more. On the other hand, occupational following leads workers into firms that only pay 0.6% more. For these two reasons—the fact that only a small minority of occupational followers are also firm followers and the fact that occupational following leads to limited differences in firm access—I do not view firms as a first order driver of occupational following and thus do not include firms in the model.

Other The model assumes that workers have complete information regarding their skills. While learning about one's skills or comparative advantage (e.g., Papageorgiou (2014)) might lead children to follow their parents, I view this as second order in relation to one's actual level of skills in the first place and thus do not include a learning component in the model. Another mechanism that might lead to occupational following could be the fact that there is uncertainty about the worker's ability from the firm's perspective (e.g., Altonji and Pierret (2001)). This could lead to the firm's use of referral networks when hiring workers, which then might lead children to enter into their parent's occupation. For example, Hensvik and Skans (2016) document the importance of referral networks and their ability to pass on information to firms about the unobserved ability of workers. However, their results hold only for professionals and associate professionals and largely do not apply for other jobs. For that reason, I view uncertainty about worker ability to not be a primary driver of occupational following, especially in blue collar occupations.

3.7 Recap of empirical results and implications for the model

I first showed the role of pre-labor market mechanisms in driving occupational persistence: both multidimensional skills and educational choice can explain a substantial share of occupational persistence, especially for white collar occupations. Second, I showed that occupational followers are characterized by differential labor market experiences: they have occupation-specific skills in their parent's occupation and are less likely to be separated from their job when they enter their parent's occupation.

The next section combines these four channels—along with a residual channel to capture whether children have an idiosyncratic preference for their parent's occupation—into a single unified model of educational and occupational choice. I then use the model to quantify the overall importance of each mechanism in driving occupational following and each mechanism's implications for model-implied total output. I also use the model for a second purpose. Section 3.4 above showed that occupational following can lead to income gains, conditional on one's multidimensional skills and education at labor market entry. I consider whether other channels still exist—such as pre-labor market investments in skills and education—that could generate even greater gains overall, especially for the children of blue collar workers. It might be that while children have a valuable advantage in their parent's occupation, if they had further developed their multidimensional skills or had made a different educational choice before they entered the labor market, then they would have experienced even greater economic success. I examine this hypothesis by considering the impacts of various counterfactuals on overall long-term class gaps in occupational attainment and income.

4 Model of educational and occupational choice

I now develop and estimate a structural model of educational and occupational choice. The model is composed of two stages: (1) an educational stage and (2) a labor market stage. In the first stage, children are 18 years old and take their set of multidimensional skills as given, which partially depend on their parental background. They then decide whether to attend college, and if so, which college major to pursue. When making the choice in the education stage, children take as given their individual-specific net value of their choice in the labor market. In the labor market stage, conditional on the educational stage, children enter the labor market and search for a job. The labor market is characterized by search frictions, which reflect the fact that workers do not immediately match with jobs, and thus both sides need to engage in costly search. Those search frictions are allowed to depend on whether or not one is in their parent's occupation. Children then match with occupations, and are potentially more productive in their parent's occupation. Moreover, children have a potential idiosyncratic preference to be in their parent's occupation. To summarize, the intergenerational influence has five potential roles in the model: (1) through the intergenerational transmission of multidimensional skills, (2) through the college and college major choice, (3) through an occupation-specific skill in one's parent's occupation, (4) through differential search frictions, and (5) through preferences.

4.1 Educational stage

To begin, children are 18 years old and are endowed with a set of multidimensional skills $\mathbf{x}$, which depends on their parental background. In particular, their skills depend on the occupation of their parents. This is similar to the model in Abbott et al. (2019), who let the cognitive and non-cognitive skills of children depend on their mother's own cognition and education. Specifically, $\mathbf{x} \in \mathbb{R}^K$ where K is the number of dimensions of skills. I use the same four skills (quantitative, verbal, social, and manual) as above. Given parental occupation

p, children draw from a skill-specific normal distribution with mean and standard deviation taken from the actual estimation sample: $x_k \sim \mathcal{N}(\mu_{kp}, \sigma_{kp}^2)$ for each skill k.

Children take their skills $\mathbf{x}$ as given and then decide whether or not to attend college, and if so, which college major to pursue. They choose a college major that they have the given skill requirements for and that sets them up for success in the labor market, foreseeing what their labor market opportunities will be given their college major choice. There are $M+1$ potential options in the set of college majors $\mathcal{M} = \{m_0, m_1, \dots, m_M\}$, which includes the option of not attending college, m_0. Children, however, face constraints in the majors they can pursue, given their skills $\mathbf{x}$.[27] Specifically, given their skills, children can choose any major m from a subset of college majors given by $\mathcal{C}(\mathbf{x}) \subseteq \mathcal{M}$. In particular, I define the college major choice set to include majors for which individuals have quantitative skills and verbal skills above their respective tenth percentile of the realized major-specific distribution of skills. I only consider the cognitive skill requirements of majors, and do not consider the social and manual skill dimensions; this closely mirrors the academic exams that lead to college degree admission, which typically focus on quantitative and verbal skills. It is important to include this constraint, so as to separate the influence of choice constraints from other factors, like preferences or differences in search and matching, on college major choice.

Children face a psychic cost of education, which I denote as $\kappa(m|\mathbf{x}, p, \nu)$ and which reflects potential knowledge and information needed to succeed in the college major, relative preparation for the major, an idiosyncratic taste for the major, and other potential factors.[28] Individuals then choose a college major to maximize the value the given major will yield in the labor market net of the cost of acquiring that major. Specifically, they solve the following optimization problem

$$\max_{m \in \mathcal{C}(\mathbf{x})} E[V(m|\mathbf{x}, p, \nu)] - \kappa(m|\mathbf{x}, p, \nu),$$

where $E[V(m|\mathbf{x}, p, \nu)]$ is the individual-specific expected value of major m in the labor market and $\kappa(m|\mathbf{x}, p, \nu)$ is the individual-specific total cost of acquiring major m. The total cost of acquiring major m is the sum of a general college cost component $\kappa^c(\mathbf{x}, p, \nu)$ and a major-

[27] This reflects the general fact that some college majors are more competitive than others. More specifically, in Finland and many countries around the world, there are minimum academic requirements to enter different college majors. For example, many countries assign students to college majors directly given a deferred acceptance algorithm that relies on standardized test scores. For instance, both Norway (Kirkeboen et al. (2016)) and Chile (Hastings et al. (2013)) require applicants to apply to both institutions and majors when applying for college. In the United States, while colleges often accept students as undeclared majors, many majors still have skill requirements (e.g., Bleemer and Mehta (2022)).

[28] See Heckman et al. (2006) for a discussion of the importance of psychic costs of education.

specific cost component $\kappa^m(\mathbf{x}, p, \nu)$:

$$\kappa(m|\mathbf{x}, p, \nu) = \kappa^c(\mathbf{x}, p, \nu) + \kappa^m(\mathbf{x}, p, \nu). \tag{4}$$

The general component, which describes the cost of college overall, is specified as

$$\kappa^c(\mathbf{x}, p, \nu) = \phi_0^c + \phi_1^c \mathbb{1}_{\text{parent's occ. requires college}} + \sum_{k=q,v,s,m} \phi_{2k}^c x_k + \phi_3^c \nu_c, \tag{5}$$

where $\mathbb{1}_{\text{parent's occ. requires college}}$ is an indicator for whether the child's parent is in an occupation that requires a college degree.[29] The child's vector of skills $\mathbf{x}$ are also allowed to affect this cost, representing general preparedness for college. Lastly, children have an idiosyncratic preference shock for college ν_c, which is drawn from a standard normal distribution.

The major-specific component is denoted as κ^m and is allowed to depend on one's parental occupation (e.g., this reflects the fact that children might be more prepared to enter medical school if their parent is a health professional, all else equal). In particular, I specify the major-specific cost as

$$\kappa^m(\mathbf{x}, p, \nu) = \phi_1^m \mathbb{1}_{\text{major's modal occ.} \neq \text{parent's occ.}} + \phi_2^m \nu_m, \tag{6}$$

where $\mathbb{1}_{\text{major's modal occ.} \neq \text{parent's occ.}}$ is an indicator for whether the modal occupation that major m leads to is different from their parent's occupation and ν_m is a major-specific preference shock drawn from a standard normal distribution.

While the educational stage is intentionally parsimonious in order to be able to combine it with a rich labor market stage in a tractable manner, it still captures the key economic forces driving educational choice. To summarize, there are various intergenerational links in the educational stage. First, parents influence the skills of their children. On one hand, this affects constraints the children face by directly affecting the college major choice set $\mathcal{C}(\mathbf{x})$ of their children. On the other hand, this affects the cost of each major through the cost function parameters ϕ_{2k}^c. Second, conditional on skills, parents can still impact the college choice and major choice of their children through the cost parameters ϕ_1^c and ϕ_1^m, respectively. Third, and lastly, parents influence the college choice through the value of each college major in the labor market, $V(m|\mathbf{x}, p, \nu)$, which I turn to next.

[29] I do not include other dimensions of parental heterogeneity, such as parental education, in the model for computational tractability. I thus use whether or not the parent is in an occupation that requires a college degree as a proxy for whether the parent has a college degree. In particular, I define this as whether the modal college choice for the given occupation is attending college or not.

4.2 Labor market stage

In the second stage of the model, children take their educational choice as given and enter the labor market. Children i are now characterized by their multidimensional skills $\mathbf{x}$, their college major choice m, and the occupation of their parents p. I model the labor market stage with a search and matching framework, largely following the model in Lise and Postel-Vinay (2020), with some key differences. There are both direct and indirect intergenerational links in the labor market stage. On one hand, parents directly influence the labor market stage by potentially affecting their children's productivity in their occupation, their children's preference for their occupation, and the search process. On the other hand, they indirectly influence the labor market stage by affecting the matching process through their effect on skills and college major choice in the educational stage.

Environment Children are characterized by their skills $\mathbf{x}$ (taken as given in the education stage as discussed above) and their chosen college major m. They then search for a job, which is defined as an occupation. An occupation j is defined by its skill requirements $\mathbf{y}^j \in \mathbb{R}^K$, as defined above in Section 2.4. When workers and jobs match, they produce output equal to $f(\mathbf{x}, \mathbf{y}, p)$, where p is the occupation of the child's parent. While in a given occupation, worker's skills accumulate, such that $\dot{\mathbf{x}} = \mathbf{h}(\mathbf{x}, \mathbf{y})$.

Workers do not immediately match with jobs and thus must engage in search. Workers can either be unemployed or employed and matched with a job. When unemployed, workers receive take-it-or-leave-it job offers from firms at the rate λ_0 from the occupational sampling distribution $\Upsilon(\mathbf{y}|m)$, which depends on the major m of the child. This specification creates the link between the educational stage and the labor market stage: the major choice in the educational stage restricts the sampling distribution of occupations in the labor market stage.[30] When employed, workers can receive outside job offers, at the rate λ_1 from the same sampling distribution $\Upsilon(\mathbf{y}|m)$. In this case, firms and workers engage in Bertrand competition and the wage is set accordingly and as described below. Workers lose their job

[30] Appendix Figure A6 illustrates the sampling distribution $\Upsilon(\mathbf{y}|m)$ by showing the probability of receiving an offer from a given occupation as a function of the college major m. For illustrative purposes, Appendix Figure A6 shows occupations collapsed from the two-digit level to the one-digit level. As shown, for example, the likelihood of receiving a job offer from a white collar occupation (one-digit codes 1, 2, and 3) is less than 20% if one does not attend college (category "None"). For most other majors, the likelihood is around 60-80%. Appendix Figure A7 illustrates this for all two-digit occupations that are in the one-digit code 2 (professionals) and 3 (associate professionals). As shown, for example, conditional on studying education, individuals have a 65% chance of receiving an offer in the teaching professions. Conditional on studying health, individuals have a 51% chance of receiving an offer to be a health professional or associate professional. On the other hand, for some majors, the link between education and occupation is less strong; for instance, studying the arts and humanities or the social sciences can lead to a vast array of occupations.

when employed at rate $\delta(\mathbf{x}, \mathbf{y}, p)$, which depends on their parental occupation p.[31]

Lastly, workers have linear preferences over income and have discount rate r. Flow utility for employed workers is given by $w - c(\mathbf{x}, \mathbf{y}, p)$, where w is the wage and $c(\mathbf{x}, \mathbf{y}, p)$ is the disutility of work, which is allowed to depend on the characteristics of the match and whether the child's occupation is the parent's occupation p. This allows for the fact that children might have idiosyncratic preferences to enter into their parent's occupation. Unemployed workers have flow utility equal to $b(\mathbf{x})$.

Value functions and wage setting Given that the search and matching structure largely follows the model in Lise and Postel-Vinay (2020), and for sake of brevity, I leave discussion of the value functions and wage-setting to Appendix Section F.

Specification I now parameterize the labor market stage of the model, and in particular, I specify the functional form for the production function $f(\cdot)$, the skill accumulation function $h(\cdot)$, the flow disutility of work $c(\cdot)$, the job separation rate $\delta(\cdot)$, and unemployment income $b(\cdot)$.

I consider four dimensions of skills: quantitative, verbal, social, and manual, as discussed in Section 2.4. Thus, the worker's skills are denoted as $\mathbf{x} = (x_q, x_v, x_s, x_m)$, while occupations are characterized by the same set of skills $\mathbf{y} = (y_q, y_v, y_s, y_m)$. Workers are also characterized in terms of a "general efficiency" skill, x_T, which helps capture the returns to experience.

I specify the production function $f(\mathbf{x}, \mathbf{y}, p)$ as follows:

$$f(\mathbf{x}, \mathbf{y}, p) = x_T \times [\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x x_k + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=p}]. \tag{7}$$

Production depends on the worker's skills and the occupational skill requirements. Furthermore, I allow for a match component between the parent's occupation and the child's occupation. Children have the potential to be more or less productive in their parent's occupation, as captured with the parameter π_1. This is motivated by the evidence that occupational following leads to greater income, as shown in Section 3.4.

While the skill requirements of occupations are constant over time, the skills of workers

[31]The only way in which I allow for intergenerational links in search is through the job separation rate. I do this to be consistent with the empirical evidence from Section 3.5: I find that occupational following reduces job separations, but has no effect on other job transitions.

accumulate over time. I specify skill accumulation as follows:

$$\mathbf{h}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} \dot{x_q} \\ \dot{x_v} \\ \dot{x_s} \\ \dot{x_m} \end{pmatrix} = \begin{pmatrix} \gamma_q|y_q - x_q| \\ \gamma_v|y_v - x_v| \\ \gamma_s|y_s - x_s| \\ \gamma_m|y_m - x_m| \end{pmatrix}. \tag{8}$$

Worker's skills thus depend on the previous period as follows: $x_k(t+1) = y_k(t) - e^{\gamma_k}(y_k(t) - x_k(t))$ for $k = q, v, s, m$ (see Appendix G.2 for the proof).

I specify the general efficiency of the worker to grow at a constant rate g and so $x_T(t) = x_T(0)e^{gt}$ (see Appendix G.2 for the proof). This can also be written as $x_T(t+1) = x_T(t)e^g$.

I specify the flow disutility of work as $c(\mathbf{x},\mathbf{y},p) = x_T \times \pi_2 \mathbb{1}_{\mathbf{y}=p}$, where π_2 represents whether the child has an idiosyncratic preference for their parent's occupation. I specify the job separation rate to depend linearly on whether the child is in their parent's occupation or not: $\delta(\mathbf{x},\mathbf{y},p) = \delta_0 + \delta_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}$. Lastly, I specify unemployment income as depending solely on the general skill: $b(\mathbf{x}) = bx_T$, with $b > 0$. This specification then gives the value of unemployment $U(x) = \frac{bx_T}{r-g}$ (see Appendix G.3 for the proof).

Now that I have specified all required model elements, I then solve for the match value $P(\mathbf{x},\mathbf{y})$ and total surplus in Appendix Section G.4. The value for an employed worker $W(\mathbf{x},\mathbf{y},\sigma)$ is then a function of known values in $U(\mathbf{x})$ and $P(\mathbf{x},\mathbf{y})$, as given in equation 23. Finally, the wage equation $w(\mathbf{x},\mathbf{y},\sigma)$ is now a function of known values, as given in equation 26.

5 Model estimation and identification

5.1 Estimation

I estimate the model in two steps. In the first step, I estimate certain parameters outside the model, including the parameters governing the intergenerational transmission of multidimensional skills, the occupational sampling distribution, and the job separation rate. I also set the discount rate in the first step. See Appendix H.1 for more details.

In the second step, I use indirect inference to estimate various parameters inside the model. These parameters include parameters related to the college and college major cost functions, the production function, the skill accumulation function, the utility function, the job offer rates, and unemployment income. I estimate parameters $\theta = [\theta_1 \ldots \theta_P]^T$ that minimize the criterion function, as given by

$$\hat{\theta} = \arg\min_{\theta} g(\theta)^T W g(\theta), \tag{9}$$

where $m_S(\theta)$ is a $P \times 1$ matrix composed of the simulated moments, m_D is a $P \times 1$ matrix composed of the data moments, W is a $P \times P$ weighting matrix, and $g(\theta)$ is a vector of percentage deviations between the simulated moments and the data moments. In particular, $g_p(\theta) = (m_{Sp}(\theta) - m_{Dp})/m_{Dp}$ for all $p = 1, \ldots, P$. Finally, I discuss the choice of the weighting matrix, how I compute the standard errors, and other estimation details in Appendix H.2.

Recall that children choose a college major in the education stage depending on the expected value they will receive from it in the labor market, denoted $E[V(m|\mathbf{x}, p, \nu)]$. In practice, I first estimate the labor market stage, which gives the labor market parameters and thus provides the expected value of each major, conditional on one's skills and parental occupation. I then estimate the education stage, conditional on the solution to the labor market stage.[32] I discuss the actual model simulation in Appendix Section H.3.

5.2 Identification

I now provide an intuitive overview of how the parameters are identified, which I also discuss in Appendix Table A22. Working backwards, I mirror the model estimation and discuss the main targeted moments for each set of parameters.[33]

Labor market stage identification The productivity of skills and occupational skill requirements are primarily identified by wages, as a function of the corresponding empirical analogues. The productivity of being in the same occupation as one's parent is primarily identified by the difference in wages between followers and non-followers.[34] As in Lise and Postel-Vinay (2020), I am inferring the productivity of the multidimensional skills, as well as one's occupation-specific skill in their parent's occupation in my case, from differences in wages. The fact that the returns to occupational following are not driven by firm access, as shown in Section 3.4, supports the interpretation that the wage premium due to occupational

[32] I am able to estimate the two stages separately because the model has no persistent unobserved heterogeneity that is common between the education stage and the labor market stage.

[33] The identification arguments in this section are further supported by the sensitivity analysis proposed in Andrews et al. (2017), which I show in Appendix Section H.6.

[34] I use the OLS estimate from the estimation sample because the 2SLS estimate from the estimation sample is much noisier than the main estimate from Table 4, due to the fact that the sample size is significantly smaller. The OLS estimate in the estimation sample is 0.023, which is in the 90% confidence interval of the main estimate from Table 4, suggesting that the difference is not especially large. If the true causal estimate was larger than 0.023, the model results would underestimate the role of occupation-specific skills.

following is productive.[35]

The sets of occupations that are acceptable to workers with the same initial skills but with varying occupational histories and parental occupations help identify the speed of human capital accumulation and the idiosyncratic preference for occupational following. To further help identify the speed of human capital accumulation, I target the correlations of individual skills with occupational skill requirements. To help identify the idiosyncratic preference for occupational following, I target the overall rate of occupational following.

In terms of the search parameters, the sample mean unemployment-to-employment transition rate helps identify the job offer rate λ_0. The mean employment-to-employment transition rate similarly helps identify the job offer rate λ_1.

Education stage identification Identification in the education stage is conditional on the individual-specific (i.e., given the skills and parental occupation of the individual) labor market values of each potential college and college major choice. The degree to which skills affect the cost of college is primarily identified by differences in college completion by individuals' skill levels. The general and differential costs of college are identified by further differences in college completion given one's multidimensional skill vector and whether one's parent is in an occupation that requires a college degree. Lastly, the major-specific college cost component is primarily identified by the share of children in college that choose a major that leads to their parent's occupation.

6 Model results and implications

I now discuss the results and implications of the model. I start by discussing the parameter estimates and model fit. I then conduct a model-based decomposition of occupational following and consider the implications of each mechanism for model-implied total output. Finally, I conclude by estimating the sources of long-term class gaps in social mobility, and their corresponding implications for policy.

[35]Nonetheless, the degree to which the occupation-specific skill represents a productive skill, rather than some other non-productive advantage, is a bound for the degree to which output is affected by these skills. Separately, the interpretation that the occupation-specific skill is productive is independent of the role of that skill in driving occupational choice in the first place; in other words, workers only value total income and would choose to enter into their parent's occupation regardless of whether their occupation-specific advantage is productive or not, as in either case it will increase their income.

6.1 Parameter estimates and model fit

Table 6 shows all the main parameter estimates, for both the education stage and the labor market stage.

Education stage parameter estimates As shown in Panel A, while there is a general cost of college (ϕ_0^c), having a parent in an occupation that requires a college degree reduces that cost by 26.6% (ϕ_1^c).[36] Furthermore, having greater cognitive skills (ϕ_{2q}^c, ϕ_{2v}^c) and social skills (ϕ_{2s}^c) reduces the cost of college as well. Again, parents have both a direct and indirect effect on the college choice: they directly reduce the cost via ϕ_1^c and they indirectly reduce the cost via their impacts on skills through ϕ_{2k}^c for $k = q, v, s, m$. The average difference in the cost of college due to differences in skills between the children of parents in an occupation that requires a college degree and an occupation that does not is about two and a half times the general cost of college overall, suggesting that the indirect intergenerational effect on college attainment is quantitatively most important.[37] Finally, the major-specific cost (ϕ_1^m) is significant, and nearly as large as the general cost of college overall.

Figure 5 shows that children face very different individual-specific eligibility as well as values and costs of each college major, by parental occupation. In Panel (a), I show that children born to white collar parents are in general more likely to have the requisite skill requirements for different college majors. For instance, consider the health major: 71.9% of the children of medical professionals have the skills required for the major, while only 53.3% of all other children do. Panel (b) shows that the net value of each major also varies significantly across the parental occupation distribution. The net value of the health major is 8.1% greater for the children of health professionals compared to all other children.[38] These initial differences in constraints and values have implications for who then becomes a health professional in the labor market. More generally, these model specifications richly capture how parents can influence educational choices—through both constraints and returns—and thus play an important pre-labor market role in shaping the potential labor market success of their children, before their children even enter the labor market.

[36] Note again that parents are only characterized by their occupation in the model. I therefore consider whether the parent is in an occupation that requires a college degree. This broadly aligns with whether or not the occupation is a white collar occupation.

[37] The average difference in skills is 0.41, 0.37, 0.19 and -0.19 standard deviations for quantitative, verbal, social, and manual skills, respectively. Multiplying these differences by the respective estimates of ϕ_{2k}^c yields 17,271, which is about two and a half times the estimate of ϕ_0^c.

[38] The net value of the major includes heterogeneity in both the gross value of the major and in the cost of the major.

Table 6: Model parameter estimates

Panel A. Education stage

General			Major
ϕ_0^c	ϕ_1^c	ϕ_3^c	ϕ_1^m
7114.73	-1892.41	35016.57	5445.53
(429.18)	(813.65)	(1976.76)	(298.73)

General			
ϕ_{2q}^c	ϕ_{2v}^c	ϕ_{2s}^c	ϕ_{2m}^c
-24935.11	-12589.28	-7630.08	4591.07
(1436.49)	(498.28)	(722.39)	(881.5)

Panel B. Labor market stage

Production function			
α_q^x	α_v^x	α_s^x	α_m^x
9708.8	147.54	1676.8	11984.63
(298.37)	(318.87)	(65.1)	(762.05)

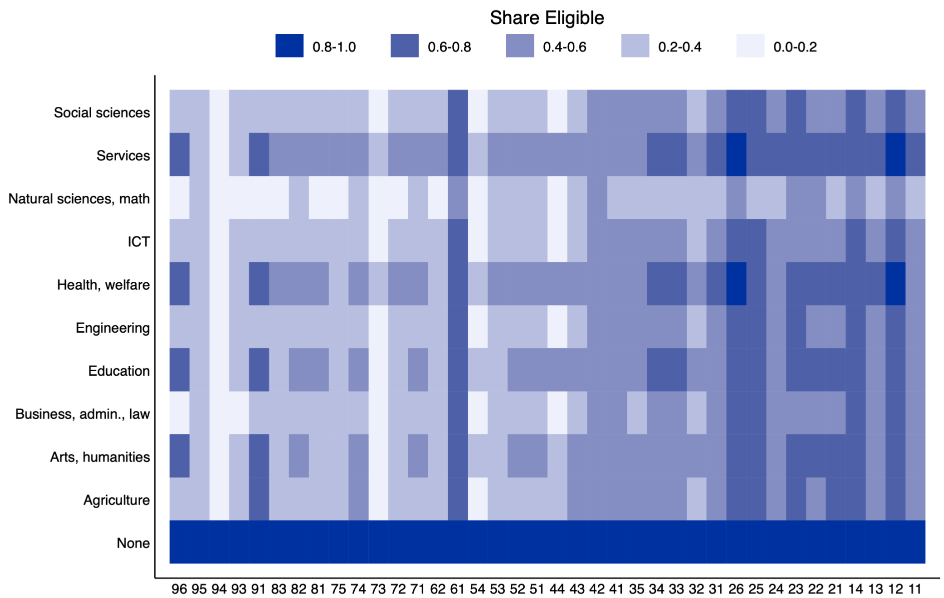
Production function				
α_q^y	α_v^y	α_s^y	α_m^y	π_1
5217.77	856.21	3090.87	340.0	498.35
(3946.66)	(509.39)	(2219.78)	(1077.68)	(31.16)

Skill accumulation function				
γ_q	γ_v	γ_s	γ_m	g
0.12	0.46	0.09	0.4	0.04
(0.0)	(0.03)	(0.0)	(0.03)	(0.0)

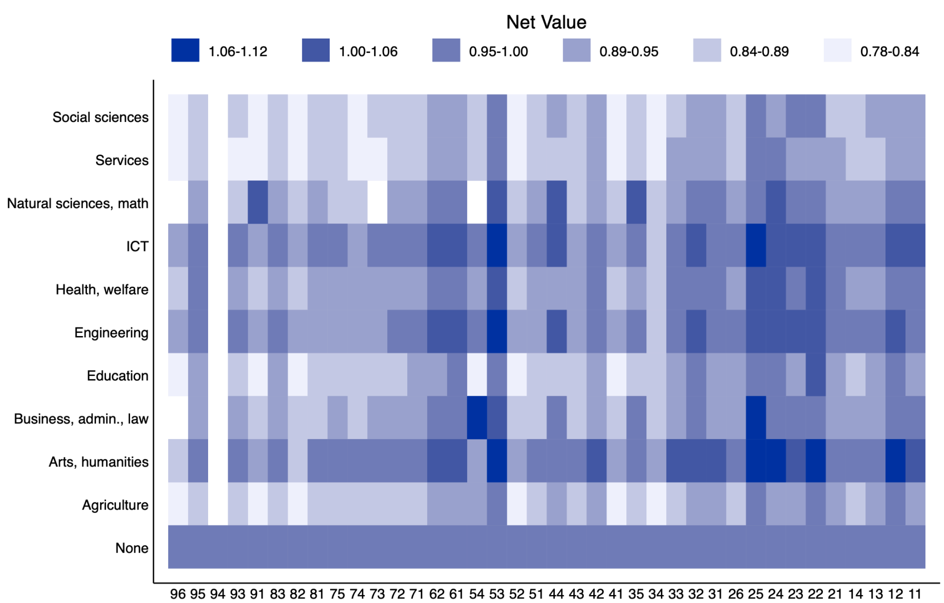
Disutility	Search		Un. Inc.
π_2	λ_0	λ_1	b
-84.09	0.25	0.85	10347.67
(62.09)	(0.0)	(0.01)	(797.16)

Notes: Parameter estimates shown. Standard errors, as described in Section 5.1, shown in parentheses.

Figure 5: Individual-specific eligibility and values of college majors



(a) Eligibility



(b) Values

Notes: Each figure shows the average for each major (on vertical axis) by parental occupation (on horizontal axis). Panel (a) shows the fraction of children that are eligible for each college major given their skills. Eligibility is defined as in Section 4.1. Panel (b) shows the average net value (gross value minus cost) of each college major, as given by the estimated parameter values and the resulting net expected value of the major. Panel (b) is limited to the children that are eligible for the given major.

Labor market stage parameter estimates As shown in Panel B of Table 6, worker skills and occupations with larger skill requirements are more productive across all skill dimensions. Quantitative and manual skills, along with quantitative skill requirements, are the most productive. Workers are also more productive in their parent's occupation, as given by π_1 and consistent with the findings from Section 3.4. The occupation-specific skill in one's parent occupation is as productive as around 5% of a standard deviation of a quantitative skill. Skills accumulate at varying speeds, which are qualitatively similar to the estimates in Lise and Postel-Vinay (2020). Social skills ($\gamma_s = 0.09$) and quantitative skills ($\gamma_s = 0.12$) adjust the slowest, while manual skills ($\gamma_m = 0.40$) and verbal skills ($\gamma_v = 0.46$) adjust faster. Next, workers have an idiosyncratic preference for their parent's occupation, although the magnitude is relatively small: in monetary terms, it is equivalent to 0.3% of average annual income.[39] Lastly, I estimate the annual job offer rate out of unemployment to be 0.25 and out of employment to be 0.85.

Model fit The model fits the data well along various targeted and non-targeted dimensions, as shown in Figure 6. A few patterns are noteworthy. In Panel (a) of Figure 6, I show that the model replicates the college attainment gradients by parental background and child skill level. Panel (b) shows that the model replicates the college major distribution, with the exception of the science and math majors. In Panel (e), I show that the wage level in the model generally aligns with the wage level in the data, but overstates it in the beginning of the life cycle. In Panel (f), while all the coefficients in the wage regression are of the expected sign, the magnitudes sometimes differ (the returns to the skills and the returns to quantitative and manual skill requirements are estimated below their true values). In Panel (h), while the model accurately reflects the rate of employment-to-employment transitions overall, it overstates it in the beginning and understates it in the end of the life cycle. Finally, Appendix Table A23 shows the data values and model values for all targeted moments.

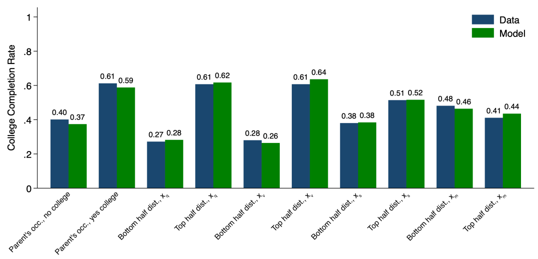
6.2 Model-based decomposition of occupational following

I now conduct a model-based decomposition of occupational following in order to quantify the role of each mechanism in driving occupational following as well as assess whether each mechanism has implications for model-implied total output. As shown in Column (3) of Table 7, the baseline rate of two-digit occupational persistence in the model is 13.8%, which is slightly larger than the rate in the data. I consider how the simulated rate of persistence

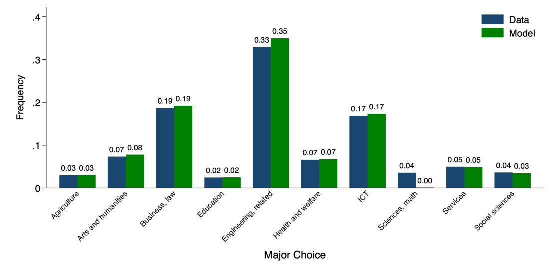
[39]This is computed by multiplying π_2 by the average general efficiency x_T and then dividing by the average annual income in the sample, which is €27,175. This can be done because total utility is specified in monetary terms.

Figure 6: Model fit

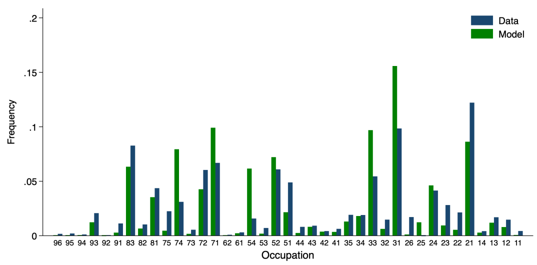
(a) College choice



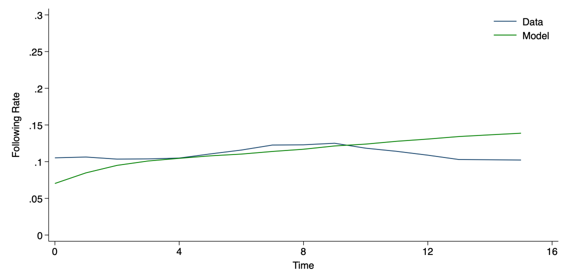
(b) Major choice



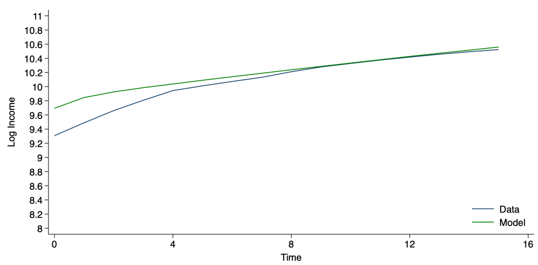
(c) Occupational distribution



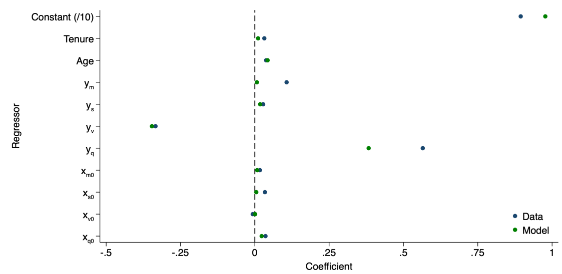
(d) Occupational following rate



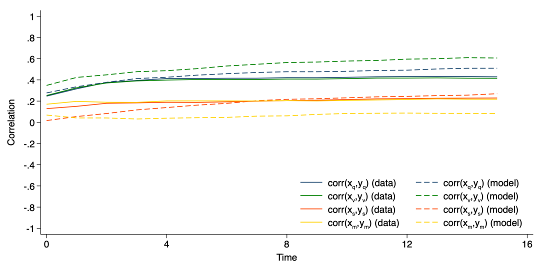
(e) Wage level



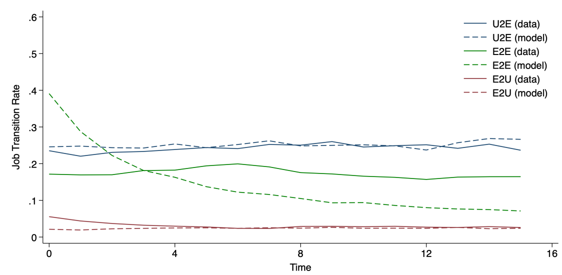
(f) Wage regression



(g) Skill correlations



(h) Job transition rates



Notes: Figures illustrate corresponding moments from the data and from the model. In Panel (a), college attainment given by whether parent's occupation requires college or not and by skill level of child (top or bottom half of distribution) for each skill. Panel (b) shows share in each major, conditional on attending college. Panel (c) shows occupational distribution. Panels (d), (e), (g), and (h) show statistics by model time period. Panel (f) is the wage regression that I target in the estimation.

changes relative to the baseline case, net of persistence that would be expected under random sorting. To start, equalizing intergenerational multidimensional skill transmission would reduce overall occupational persistence by 9.0%. To be precise, if children drew from the same distribution of initial multidimensional skills, rather than a distribution that depended on parental occupation, the overall rate of persistence would decrease by 9.0%, from 13.8% to 12.8%.[40]

Table 7: Model-based decomposition of occupational persistence

	Rate of following			Total output
	Class	1-Digit	2-Digit	
	(1)	(2)	(3)	(4)
Baseline	0.537	0.246	0.138	
Equalize skill transmission	-99.2%	-16.5%	-9.0%	0.1%
	[0.5]	[0.224]	[0.128]	
Equalize college cost	-0.6%	-0.1%	0.1%	-0.0%
	[0.537]	[0.246]	[0.138]	
No major cost	-0.9%	-1.4%	-2.8%	0.0%
	[0.537]	[0.244]	[0.135]	
No idiosyncratic preferences	-12.0%	-6.5%	-8.5%	-0.01%
	[0.532]	[0.237]	[0.128]	
No differential search frictions	-1.2%	-0.7%	-1.0%	-0.03%
	[0.536]	[0.245]	[0.137]	
No occupation-specific skill	-26.4%	-29.7%	-44.4%	-0.25%
	[0.527]	[0.206]	[0.088]	
Equalize education stage	-100.9%	-17.6%	-12.6%	0.12%
	[0.5]	[0.222]	[0.124]	
Equalize labor stage	-31.4%	-37.5%	-57.7%	-0.23%
	[0.525]	[0.195]	[0.073]	
Equalize all	-130.8%	-51.2%	-64.3%	-0.07%
	[0.489]	[0.177]	[0.066]	

Notes: Percentage change in rate of occupational following from model-implied baseline case shown, with overall level shown in brackets. Percentage change is relative to random sorting benchmark, as defined in Appendix Section H.4. I compute the percentage change as $1 - \frac{\text{simulated rate} - \text{random benchmark}}{\text{baseline rate} - \text{random benchmark}}$. Occupation is taken at age 30. Shown for all children. Shown at class level, 1-digit level, and 2-digit level. Percentage change in total output from model-implied baseline case shown. All mechanisms considered on their own in isolation.

Equalizing the differential general college cost—by setting ϕ_1^c equal to zero in the general psychic cost of college (Equation 5)—would have no direct effect on occupational persistence. Removing the major-specific cost of college—by setting ϕ_1^m equal to zero in the major-specific

[40]See Appendix Section H.5 for an overview of the precise manner in which I conduct each model counterfactual.

psychic cost of college (Equation 6)—would reduce occupational following by 2.8%. Overall, equalizing the education stage of the model would reduce occupational following by 12.6%. This share (12.6%) is similar to the share (19.2%) estimated in the statistical decomposition in Sections 3.2 and 3.3 that represented the role of skills and education in driving occupational persistence.[41] The fact that these two different methods ascribe similar shares of occupational persistence to skills and education bolsters confidence in their validity.

In terms of the labor market mechanisms, differential search frictions explain only a very small share of the persistence, while idiosyncratic preferences explain a substantial share of the persistence. Equalizing differential search frictions—by setting the job separation rate to be constant across all individuals, regardless of whether they enter their parent's occupation or not—would reduce the level of following by 1.0%. Removing idiosyncratic preferences—by setting π_2 equal to zero in the specification of flow utility (Section 4.2)—would reduce the level of following by 8.5%. Lastly, occupation-specific skills explain the largest share of two-digit occupational persistence: removing them—by setting π_1 equal to zero in the specification of production (Equation 7)—would reduce the level of occupational persistence by 44.4%. Overall, the model captures the main relevant mechanisms well. All the mechanisms in the model can jointly explain 64.3% of the non-random component of two-digit occupational persistence.[42]

In terms of whether these mechanisms are productive or not, I largely find that these mechanisms do not significantly affect model-implied overall output, with the exception of the occupation-specific skills. The existence of occupation-specific skills increases model-implied overall output: they are responsible for 0.3% of total output, implying that the intergenerational transmission of occupation-specific skills is not only a key driver of occupational persistence, but also is productive in nature.[43]

Table 7 also shows the decomposition at coarser levels of occupational classification. A key difference emerges, which again is consistent with the results from Sections 3.2 and 3.3: pre-labor market differences can explain more persistence at coarser levels of occupational classification than at finer levels. For instance, equalizing the education stage would reduce

[41] The respective shares attributed to skills and to education are similar in the estimated model and in the statistical decomposition in Sections 3.2 and 3.3, although the relative importance shifts towards skills and away from education in the model-based decomposition. This is consistent with the findings from Bolt et al. (2024), who find that the importance of education in explaining the intergenerational elasticity of earnings (IGE) decreases when allowing for indirect effects of skills on the IGE through education.

[42] The remaining amount of occupational persistence is due to the random shocks in the model. This could be due to various other mechanisms that drive occupational persistence that I do not consider in the model, as discussed in Section 3.6.

[43] The relevant counterfactual is to remove occupation-specific skills, rather than redistribute them, because they cannot be simply redistributed. The occupation-specific skills represent skills that children gain from their parents directly, and thus could not be simply redistributed by any reasonable policy.

persistence entirely at the class level, compared to by only 12.6% at the two-digit level. On the other hand, the importance of occupation-specific skills decreases when looking at coarser levels of classification. This is because occupation-specific skills shift workers between relatively similar occupations, while the more general pre-labor market skills shift workers amongst relatively more distinct occupations with larger differences in overall skill requirements.

Robustness In Appendix Table A24, I consider an alternative version of the decomposition in which I jointly consider each mechanism, in the order that parallels the chronological nature of the model, rather than considering each mechanism in isolation on its own. While I find that the results are extremely similar, the importance of each channel is not additive: for example, equalizing the college cost on its own does not reduce persistence as shown in Table 7, but conditional on equalizing skill transmission, it reduces persistence by another 0.1%. This underscores the importance of considering each mechanism jointly: considering them separately can understate or overstate the relative importance of each mechanism by ignoring how they interact with one another.

6.3 Class gaps in social mobility and policy implications

With the relative importance of each mechanism in mind, I now turn to the results' implications for social mobility more broadly. While I have shown that occupational following can lead to income gains conditional on pre-labor market skills and education, I now consider whether policy levers still exist that would further increase both occupational attainment and income for individuals. I operationalize this by first decomposing the sources of class gaps in both occupational and income mobility and second simulating the impacts of actual policies. As motivated in the introduction, I consider entry into both white collar occupations overall and elite white collar occupations in particular. My approach is in contrast to recent studies on intergenerational occupational mobility that are limited in their ability to conduct policy-relevant counterfactuals. On one hand, recent work that is not based on a structural model (e.g., Ventura (2023)) is unable to conduct counterfactual analysis in the first place. On the other hand, recent model-based research (e.g., Lo Bello and Morchio (2022)) considers broad mechanisms; the model I offer here is micro-founded, therefore letting me identify precisely the relevant economic mechanisms at play and thus offer specific policy implications.

Class gaps in social mobility Table 8 illustrates the degree to which children are able to attain both white collar occupations in general and elite white collar occupations in particular. In the baseline, the overall level of white collar occupational attainment is 50.8%. However, the children of white collar parents are 16.1% more likely to become white collar than are the children of blue collar parents—this is the estimate I refer to as the class gap in white collar occupational attainment. I consider to what degree each mechanism of the model explains this class gap. First, if all children drew from the same distribution of skills, rather than a distribution segmented by parental occupation, the overall attainment of white collar occupations would not change. However, the share of blue collar children that become white collar would increase while the share of white collar children that become white collar would decrease. Taken together, this would significantly reduce the class gap, by 88%. This estimate is slightly larger than the most related recent estimates of the drivers of labor market inequality, although using different measures of inequality. Taber and Vejlin (2020) find that 59 to 82% of overall wage inequality in Denmark is due to variation in pre-labor market skills, while Lise and Postel-Vinay (2020) find that 65% of the variation in the social value of output in the United States is due to variation in pre-labor market skills.[44]

The direct intergenerational influence on college attainment, conditional on skills, is small and barely contributes to the class gap in social mobility. This implies that despite the difference in college attainment by parental background, parents do not exert a large direct influence on college attainment. Instead, the influence primarily works indirectly through differences in initial skills.[45] Occupation-specific skills are also a substantial driver of the class gap in white collar occupational attainment. However, in line with multi-dimensional skills driving class occupational persistence more than occupation-specific skills do, multi-dimensional skills drive the class gap in white collar attainment more than occupation-specific skills do. Idiosyncratic preferences and differential search frictions contribute very little to the class gap.

I define elite occupations at the two-digit level and consider an index composed of the health, business, and legal professions as they are often seen as elite occupations in society and have been a significant focus of both research and policy (as discussed in the introduction).[46] As shown in Column (2) of Table 8, equalizing skill transmission would also significantly reduce the class gap in elite occupational attainment (by 16%), although to a significantly smaller degree than their reduction in the white collar class gap.[47] On the

[44]Keane and Wolpin (1997) find that 90% of the variation in lifetime utility is due to initial skills.

[45]These results are again similar to the results from Bolt et al. (2024), who find that when accounting for the relationship between child skills and educational choice, parental earnings do not have a large impact on child earnings through educational choice.

[46]I consider two-digit occupations 22 (health professionals), 24 (business professionals), and 26 (legal and social professionals).

[47]Consistent with the fact that skills explain a smaller share of occupational persistence at the two digit-level than at the class level, they explain a smaller share of the class gap in elite occupational attainment

Table 8: Sources of long-term class gaps in social mobility

	Class gap		
	WC	Elite	Income
	(1)	(2)	(3)
Baseline	0.161	0.155	0.031
Equalize skill transmission	-88.3%	-15.8%	-98.6%
	[0.019]	[0.131]	[0.0]
Equalize college cost	-0.7%	-5.2%	4.8%
	[0.16]	[0.147]	[0.033]
No major cost	0.2%	-32.8%	-0.7%
	[0.161]	[0.104]	[0.031]
No idiosyncratic preferences	-6.9%	-9.1%	-2.7%
	[0.15]	[0.141]	[0.031]
No differential search frictions	-0.3%	0.0%	-0.6%
	[0.16]	[0.155]	[0.031]
No occupation-specific skill	-17.5%	-20.6%	-10.7%
	[0.133]	[0.123]	[0.028]
Equalize education stage	-86.6%	-41.9%	-100.2%
	[0.022]	[0.09]	[-0.0]
Equalize labor stage	-19.0%	-12.9%	-10.8%
	[0.13]	[0.135]	[0.028]
Equalize all	-102.4%	-38.8%	-109.0%
	[-0.004]	[0.095]	[-0.003]

Notes: Column (1) shows relative difference in likelihood of attaining white collar occupation by parental class. Column (2) shows relative difference in likelihood of attaining elite occupation (medicine, business, or law; two-digit occupations 22, 24, and 26) by parental class. Column (3) shows relative difference in income rank by parental class. Occupation and income are taken at age 30. For all columns, change from model-implied baseline case shown, with class gap shown in brackets.

other hand, occupation-specific skills are more influential in driving the elite occupational attainment class gap. The most significant driver of the elite occupational attainment class gap are the major-specific costs of college, which include information frictions, academic preparation for different college majors, and costs of foregone earnings, for example. These major-specific frictions in general are responsible for nearly one third (33%) of the elite class gap. This suggests that policies that reduces frictions into certain majors, such as medicine (e.g., Jiménez (2024)), would be effective in increase mobility into elite occupations.

Lastly, Column (3) of Table 8 shows the key sources of the class gap in income—defined by the relative difference in expected income rank by parental class. Consistent with the key drivers of occupational mobility, differences in initial skill endowments are responsible for almost the entirety of the class gap in income. While significantly less important than the pre-labor market skills, occupation-specific skills are responsible for 11% of the class gap in income.

Implications for policy The decomposition exercises above suggest that—despite the occupation-specific skills children have in their parent's occupation—pre-labor market skills and education are still relevant policy levers to further increase social mobility. First, pre-labor market skills (i.e., the quantitative, verbal, social, and manual skills of individuals) are a key driver of long-run class gaps in occupational attainment. They can explain 88% and 16% of the class gap in white collar occupational attainment and elite occupational attainment, respectively. Second, occupation-specific skills are also important, and can explain 18% and 21% of the respective gaps. Third, major-specific frictions in educational choice are the most important drivers behind the class gap in elite occupational attainment. Overall, this suggests that two classes of policies would be effective in increasing mobility: (1) policies that invest in early-life skill development and cultivate a broad set of skills in children (e.g., policies that promote early childhood development; Duncan et al. (2023)) and (2) policies that make high-return college majors more accessible (e.g., information provisions; Hastings et al. (2016)). Furthermore, the decomposition results imply that such investments in pre-labor market skills and educational choice would not come at a cost to overall output.

I now consider how actual policies can be used to impact occupational attainment and income. Given the importance of multidimensional skills in driving occupational persistence and class gaps in social mobility, I consider policies that target skill development. First, a historical school reform in Finland provides a useful case study. In the 1970s, Finland implemented a wide-reaching school reform, which included delaying the age at which chil-

(which is defined at the two-digit level) than the class gap in white collar occupational attainment (which is defined at the class level).

dren would track into different school paths from age 11 to age 15. This policy had no net effect on the cognitive skills of children (as measured by the same measures from the Finnish Defence Forces that I use), but a significant distributional effect: it increased the skills of children of low-educated parents by 0.05 standard deviations, while decreasing the skills of the children of highly-educated parents by 0.04 standard deviations (Kerr et al. (2013)). Overall, this reduced the gap in cognitive skills by parental background by 0.08 standard deviations, or 29.8%. These results are an example of how the structure of an education system can have meaningful impacts on skill development, and also serve as evidence that the skills measured by the Defence Forces are malleable (as opposed to being fixed from a young age). I then calculate the model-implied impacts of a reduction in the class gap of cognitive skills by 29.8%, mirroring the impacts found in Kerr et al. (2013). As shown in Table 9, this reduction in the differences of initial skill endowments implies a reduction in the white collar class gap by nearly one fourth, in addition to significant reductions in the elite class gap and income class gap.

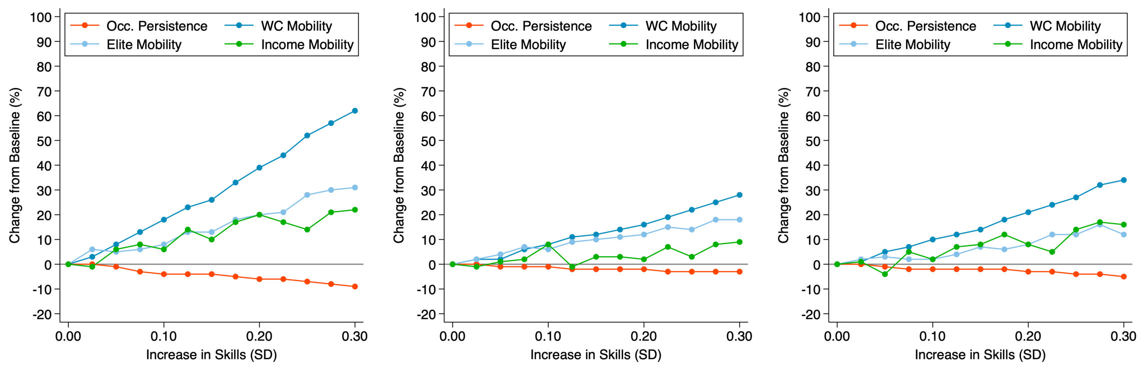
Table 9: Policy application I: Simulated impact of an education system reform on long-term class gaps in social mobility

	Class gap		
	Baseline (1)	Policy (2)	Change (3)
White collar occupational attainment	0.161	0.119	-23.7%
Elite occupational attainment	0.155	0.137	-11.9%
Income	0.031	0.028	-8.3%

Notes: White collar occupational attainment class gap is defined by difference in likelihood of becoming white collar by parental class. Elite occupational attainment class gap is defined by difference in likelihood of becoming elite (entering 2-digit occupation 22, 24, or 26) by parental class. Income class gap is defined by difference in average income rank by parental class. Occupation and income are taken at age 30. Column (1) shows baseline case, Column (2) shows results under policy counterfactual, and Column (3) shows the change in class gaps between Columns (1) and (2).

Second, I consider a hypothetical policy that would target the skill development of the children of blue collar workers (e.g., tutoring programs (Nickow et al. (2020)) and schooling investments in disadvantaged areas (Johnson and Jackson (2019))). In Panel (a) of Figure 7, I show that a relatively modest increase in the cognitive skills of the children of blue collar workers would lead to substantial decreases in occupational persistence and increases in social mobility. Furthermore, as shown in Panels (b) and (c), these impacts are driven by increases in both quantitative and verbal skills, rather than just either cognitive skill alone, underscoring the multidimensional nature of skills. Taken together, these two applications to policy highlight the importance of early-life differences in multidimensional skills for long-term outcomes in both occupational attainment and income.

Figure 7: Policy application II: Simulated impacts of increase of skills on social mobility



(a) Both cognitive skills (b) Quantitative skills (c) Verbal skills

Notes: Figures show change in long-term outcomes relative to baseline case due to simulated increase in cognitive skills (quantitative and verbal sklls). First, figures show change in occupational persistence relative to baseline case. Second, figures also show change in social mobility relative to baseline case. In particular, change in social mobility is defined as reduction in class gaps. Skills are measured in standard deviations (SD). Occupational persistence defined at two-digit level. Occupation and income are taken at age 30. Measures of social mobility as defined in Table 8.

Robustness to general equilibrium effects The model is set in partial equilibrium in the sense that while wages are determined via negotiation between workers and jobs as in Postel-Vinay and Robin (2002), the productivity of skills and skill requirements are held constant in all counterfactual analyses. The returns to skills and skill requirements, and thus wages, do not adjust to the overall stock of skills or education and are not set in general equilibrium. Moreover, I do not consider capacity constraints in either the number of individuals that can attend college and enter each college major as well as in the different occupations. These two assumptions, however, are both common assumptions in search and matching models (e.g., Lise and Postel-Vinay (2020) and Lindenlaub and Postel-Vinay (2023), respectively), as they greatly help with computational tractability.

These assumptions would affect the counterfactual analysis to varying degrees. On one hand, it is unlikely that these concerns would significantly affect the counterfactual of equal skill transmission, no idiosyncratic preferences, and no differential search frictions. This is because the total stock of skills, as well as the total labor supply in white collar occupations, does not significantly change under these counterfactuals, thus providing limited scope for general equilibrium effects. On the other hand, removing college or major costs has more potential to generate an equilibrium response as it would change the overall labor supply in white collar occupations and elite occupations. Despite this, it is unlikely that general equilibrium effects would entirely account for the reduction in the class gaps that I estimate. As a point of comparison, Abbott et al. (2019) considers three alternative expansions of federal student aid policies for college attainment in the U.S., and their corresponding partial equilibrium and general equilibrium impacts. In my supplementary calculations, the reduction in the class gap in college attainment by parental wealth attenuates by 16-58% when accounting for the general equilibrium response (see Appendix Section I.1 for more). This suggests that while general equilibrium effects would probably attenuate to some degree the impacts of reducing the costs of education on class gaps in occupational attainment, they are unlikely to explain most of the impacts.

7 Conclusion

The intergenerational persistence of occupation, and economic success more generally, has been recognized as an important economic phenomenon—with sizable implications for social mobility—for decades (e.g., Blau and Duncan (1967); Chetty et al. (2014); Laband and Lentz (1983); Stansbury and Rodriguez (2024)). I bring new data and use a mix of empirical methods to first rigorously characterize occupational persistence and second to identify the core microeconomic drivers of this persistence. Third, I develop and estimate a rich and novel model of educational and occupational choice to assess the relative importance of each mechanism and the aggregate consequences of occupational persistence for the economy. Accordingly, I simulate various counterfactuals to shed light on what types of policies would be most effective in reducing persistence and increasing mobility, and whether these policy impacts would affect overall output.

I find that there are sustained intergenerational links throughout the life cycle that lead to occupational persistence: from early-life skill formation, to educational choice, and to differential labor market returns and search frictions, individuals differ by parental background. I find that the presence of occupation-specific skills is both a key driver of occupational persistence and a productive one. Conditional on pre-labor market skills and education, occupational following leads to sustained increases in income, including a 5.5% increase in the first year upon labor market entry. The occupational following wage premium—in the model-based decomposition—is responsible for nearly half (44%) of occupational persistence. Moreover, the occupation-specific skills are responsible for 0.3% of model-implied total output. In line with related research that has focused on specific occupations (e.g., Laband and Lentz (1992); Ventura (2023)), this means that occupation-specific skills are a key and productive driver of persistence across all occupations on average. In terms of the other labor market mechanisms, I find that idiosyncratic preferences and differential search frictions play a small role in driving occupational persistence.

At the same time, I find that pre-labor market skills and educational choice also substantially drive persistence, and have no differential impacts on overall output. Using two different methods, I show that these pre-labor market differences can explain 13-19% of occupational persistence at the two-digit level, and the vast majority at the class level. Correspondingly, these early-life differences in skills and in education explain 87% and 42% of long-term class gaps in white collar occupational attainment and elite occupational attainment, respectively. In terms of educational choice, I find that major-specific frictions, rather than frictions associated with just attending college or not, are the most important factor in driving the class gap in elite occupational attainment. Taken together, this suggests that policies that invest in a broad set of early-life skills and in making high-return college majors more accessible would be effective in reducing long-term class gaps in economic success and thus increasing social mobility.

These results should help focus the attention of researchers and policymakers on early-life differences in explaining long-term levels of social mobility. In addition to the robust and growing body of research on the drivers of skill formation and educational choice, further investigation into these early-life factors—as well as related policies—and in particular how they relate to long-term occupational attainment and social mobility is a ripe area for future research.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante, "Education policy and intergenerational transfers in equilibrium," *Journal of Political Economy*, 2019, *127* (6), 2569–2624.

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

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

Altmejd, Adam, "Inheritance of fields of study," 2022.

Altonji, Joseph G. and Charles R. Pierret, "Employer learning and statistical discrimination," *Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

_ , Erica Blom, and Costas Meghir, "Heterogeneity in human capital investments: High school curriculum, college major, and careers," *Annual Review of Economics*, 2012, *4*, 185–223.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro, "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 2017, *132* (4), 1553–1592.

Anstreicher, Garrett, "Spatial Influences in Upward Mobility," *Journal of Political Economy*, 2024, *132* (1), 149–199.

Arellano-Bover, Jaime, "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," *Journal of Labor Economics*, 2024, *42* (2), 549–589.

Bandiera, Oriana, Ananya Kotia, Ilse Lindenlaub, Christian Moser, and Andrea Prat, "Meritocracy Across Countries," 2024.

Bertrand, Marianne, Sandra E. Black, Sissel Jensen, and Adriana Lleras-Muney, "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway," *Review of Economic Studies*, 2019, *86* (1), 191–239.

Bjorklund, Anders and Markus Jantti, "Intergenerational mobility of socio-economic status in comparative perspective," *Social Policy Research Centre*, 2000.

Black, Sandra E., Paul J. Devereux, Fanny Landaud, and Kjell G. Salvanes, "Where Does Wealth Come From? Measuring Lifetime Resources in Norway," *Journal of Economic Perspectives*, 2023, *37* (4), 115–136.

Blau, Peter M. and Otis Dudley Duncan, *The American Occupational Structure* 1967.

Bleemer, Zachary and Aashish Sunil Mehta, "College Major Restrictions and Student Stratification," *SSRN Electronic Journal*, 2022, (March).

BMA, "Widening participation in medicine," *British Medical Association*, 2023.

Boar, Corina and Dania Lashkari, "Occupational Choice and the Intergenerational Mobility of Welfare," *NBER Working Paper Series*, 2021.

Bolt, Uta, Eric French, Jamie Hentall-MacCuish, and Cormac O'Dea, "The intergenerational elasticity of earnings: Exploring the mechanisms," 2024.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ **and** _ , "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *Quarterly Journal of Economics*, 2018, *133* (3), 1163–1228.

_ **, David Deming, and John N. Friedman**, "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges," *NBER Working Paper Series*, 2023.

_ **, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez**, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Corak, Miles and Patrizio Piraino, "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 2011, *29* (1), 37–68.

Dobbin, Caue and Tom Zohar, "Quantifying the Role of Firms in Intergenerational Mobility," 2024.

Duffie, Darrell and Kenneth J. Singleton, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 1993, *61* (4), 929–952.

Duncan, Greg, Ariel Kalil, Magne Mogstad, and Mari Rege, "Investing in early childhood development in preschool and at home," *Handbook of the Economics of Education*, 2023, *6*, 1–91.

Dunn, Thomas and Douglas Holtz-Eakin, "Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links," *Journal of Labor Economics*, 2000, *18* (2), 282–305.

Emran, M. Shahe and Forhad Shilpi, "Intergenerational Occupational Mobility in Rural Economy: Evidence from Nepal and Vietnam," *Journal of Human Resources*, 2011, *46* (2), 427–458.

Engzell, Per and Nathan Wilmers, "Firms and the Intergenerational Transmission of Labor Market Advantage," 2024.

Flinn, Christopher and Joseph Mullins, "Labor market search and schooling investment," *International Economic Review*, 2015, *56* (2), 359–398.

Fredriksson, Peter, Lena Hensvik, and Oskar Nordström Skans, "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility," *American Economic Review*, 2018, *108* (11), 3303–3338.

George, Siddharth Eapen, "Like Father, Like Son? The Effect of Political Dynasties on Economic Development," 2019.

Gourieroux, C, A Monfort, and E Renault, "Indirect Inference," *Journal of Applied Econometrics*, 1993, *8*, S85–S118.

Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer, "Multidimensional skill mismatch," *American Economic Journal: Macroeconomics*, 2020, *12* (1), 210–244.

_ **, Fatih Karahan, Serdar Ozkan, and Jae Song**, "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?," *Econometrica*, 2021, *89* (5), 2303–2339.

Haeck, Catherine and Jean-William Laliberté, "Careers and Intergenerational Income Mobility," 2024.

Hastings, Justine S., Christopher A. Neilson, Anely Ramirez, and Seth Zimmerman, "(Un)informed college and major choice: Evidence from linked survey and administrative data," *Economics of Education Review*, 2016, *51*, 136–151.

— , **Christopher Neilson, and Seth Zimmerman**, "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile," *NBER Working Paper*, 2013, pp. 1–50.

Heckman, James and Rasmus Landersø, "Lessons for Americans from Denmark about inequality and social mobility," *Labour Economics*, 2022, *77* (May 2021), 101999.

Heckman, James J., Lance J. Lochner, and Petra E. Todd, "Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 2006, *1*, 307–458.

Hensvik, Lena and Oskar Nordström Skans, "Social networks, employee selection, and labor market outcomes," *Journal of Labor Economics*, 2016, *34* (4), 825–867.

Humphries, John Eric, "The Causes and Consequences of Self-Employment over the Life Cycle," 2022.

International Labour Organization, "The International Standard Classification of Occupations (ISCO-08) companion guide," 2023, (September).

Jiménez, Jesus, "$1 Billion Bloomberg Gift to Hopkins Makes Tuition Free for Most Medical Students," *The New York Times*, 2024.

Johnson, Rucker C. and C. Kirabo Jackson, "Reducing inequality through dynamic complementarity: Evidence from head start and public school spending," *American Economic Journal: Economic Policy*, 2019, *11* (4), 310–349.

Jokela, Markus, Tuomas Pekkarinen, Matti Sarvimäki, Marko Terviö, and Roope Uusitalo, "Secular rise in economically valuable personality traits," *PNAS*, 2017, *114* (25), 6527–6532.

Kaila, Martti, Emily Nix, and Krista Riukula, "The Impact of an Early Career Shock on Intergenerational Mobility," *Journal of Labor Economics*, 2024.

Kambourov, Gueorgui and Lourii Manovskii, "Occupational Specificity of Human Capital," *International Economic Review*, 2009, *50* (1), 63–115.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Kerr, Sari Pekkala, Tuomas Pekkarinen, and Roope Uusitalo, "School tracking and development of cognitive skills," *Journal of Labor Economics*, 2013, *31* (3), 577–602.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

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, *50* (2), 474.

_ and _ , "Favorite Sons: Intergenerational Wealth Transfers Among Politicians," *Economic Inquiry*, 1985, *23* (3), 395–414.

_ and _ , "Self-Recruitment in the Legal Profession," *Journal of Labor Economics*, 1992, *10* (2), 182–201.

Landersø, Rasmus and James J. Heckman, "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US," *Scandinavian Journal of Economics*, 2017, *119* (1), 178–230.

Lentz, Bernard F. and David N. Laband, "Why So Many Children of Doctors Become Doctors: Nepotism vs. Human Capital Transfers," *The Journal of Human Resources*, 1989, *24* (3), 396.

_ and _ , "Entrepreneurial Success and Occupational Inheritance among Proprietors," *The Canadian Journal of Economics*, 1990, *23* (3), 563–579.

Lindenlaub, Ilse, "Sorting Multidimensional Types: Theory and Application," *The Review of Economic Studies*, 2017, *84* (2), 718–789.

_ and Fabien Postel-Vinay, "Multidimensional Sorting Under Random Search," *Journal of Political Economy*, 2023.

Lindquist, Matthew J., Joeri Sol, and Mirjam Van Praag, "Why do entrepreneurial parents have entrepreneurial children?," *Journal of Labor Economics*, 2015, *33* (2), 269–296.

Lise, Jeremy and Fabien Postel-Vinay, "Multideimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 2020, *110* (8), 2328–2376.

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

Long, Jason and Joseph Ferrie, "Intergenerational Occupational Mobility in Great Britain and the United States Since 1850," *American Economic Review*, 2013, *103* (4), 1109–1137.

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*, 2022, *57* (2), 525–554.

Montonen, Jerry and Samuel M. Solomon, "The Impact of Peers on Occupational Choice and Long-term Outcomes," 2024.

Nickow, Andre, Philip Oreopoulos, and Vincent Quan, "The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence," *NBER Working Paper*, 2020.

OECD, "Education attainment," 2024.

Pakes, Ariel and David Pollard, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 1989, *57* (5), 1027–1057.

Papageorgiou, Theodore, "Learning your comparative advantages," *Review of Economic Studies*, 2014, *81* (3), 1263–1295.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

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 and Organization*, 2021, *184*, 489–505.

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

Staiger, Matthew, "The Intergenerational Transmission of Employers and the Earnings of Young Workers," 2023.

Stansbury, Anna and Kyra Rodriguez, "The Class Gap in Career Progression: Evidence from US academia," 2024.

Statistics Finland, "Employment," 2024.

Taber, Christopher and Rune Vejlin, "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 2020, *88* (3), 1031–1069.

Todd, Petra E. and Weilong Zhang, "A dynamic model of personality, schooling, and occupational choice," *Quantitative Economics*, 2020, *11* (1), 231–275.

U.S. Bureau of Labor Statistics, "Labor Force Statistics from the Current Population Survey," 2024.

Ventura, Maria, "Following in the family footsteps: Incidence and returns of occupational persistence," 2023.

YLE, "Finland faces public sector doctor shortage," *YLE News*, 2021.

Zimmerman, Seth, "Elite colleges and upward mobility to top jobs and top incomes," *American Economic Review*, 2019, *109* (1), 1–47.

A Appendix Tables

Table A1: Occupational following rates at different levels of classification

	Labor market entry			Age 35		
	All	Males	Females	All	Males	Females
Either parent						
Same 1-digit occupation	0.331	0.317	0.344	0.344	0.337	0.352
Same 2-digit occupation	0.132	0.128	0.135	0.133	0.132	0.134
Same 3-digit occupation	0.073	0.074	0.071	0.074	0.075	0.072
Same 4-digit occupation	0.053	0.051	0.054	0.052	0.053	0.051
Ever same 1-digit occupation				0.646	0.641	0.651
Ever same 2-digit occupation				0.331	0.332	0.329
Ever same 3-digit occupation				0.201	0.208	0.195
Ever same 4-digit occupation				0.149	0.147	0.152
Father						
Same 1-digit occupation	0.198	0.224	0.174	0.212	0.241	0.180
Same 2-digit occupation	0.075	0.102	0.049	0.079	0.108	0.048
Same 3-digit occupation	0.041	0.062	0.020	0.044	0.065	0.021
Same 4-digit occupation	0.026	0.042	0.013	0.031	0.048	0.013
Ever same 1-digit occupation				0.442	0.499	0.380
Ever same 2-digit occupation				0.206	0.269	0.138
Ever same 3-digit occupation				0.120	0.174	0.062
Ever same 4-digit occupation				0.083	0.121	0.042
Mother						
Same 1-digit occupation	0.217	0.172	0.260	0.219	0.179	0.262
Same 2-digit occupation	0.077	0.047	0.106	0.075	0.046	0.105
Same 3-digit occupation	0.042	0.024	0.060	0.041	0.023	0.059
Same 4-digit occupation	0.035	0.020	0.048	0.030	0.017	0.044
Ever same 1-digit occupation				0.484	0.428	0.544
Ever same 2-digit occupation				0.200	0.144	0.260
Ever same 3-digit occupation				0.116	0.075	0.159
Ever same 4-digit occupation				0.093	0.057	0.130
Observations	760344	372567	387777	789505	408066	381439

Notes: Overall population considered. Sample means shown. Shown with respect to parent indicated.

Table A2: Correlation between skills

	Quant.	Verbal	Social	Manual
Quant.	1			
Verbal	0.68	1		
Social	0.12	0.13	1	
Manual	-0.03	-0.08	0.06	1

Notes: Table shows pairwise correlation between all skills in main estimation sample.

Table A3: Predictive power of skill measures: occupational choices

	x_q	x_v	x_s	x_m
1. Managers	0.003***	0.001	0.005***	-0.002***
(Mean: 0.013)	(0.001)	(0.001)	(0.000)	(0.000)
2. Professionals	0.054***	0.050***	0.010***	-0.037***
(Mean: 0.183)	(0.002)	(0.002)	(0.001)	(0.001)
3. Assoc. professionals	0.045***	0.034***	0.015***	-0.004***
(Mean: 0.217)	(0.002)	(0.002)	(0.001)	(0.001)
4. Clerical support	0.008***	0.006***	0.002***	-0.007***
(Mean: 0.045)	(0.001)	(0.001)	(0.001)	(0.001)
5. Service, sales	-0.020***	-0.001	0.021***	-0.007***
(Mean: 0.133)	(0.002)	(0.002)	(0.001)	(0.001)
6. Agriculture	-0.003***	-0.001***	-0.002***	0.001***
(Mean: 0.006)	(0.000)	(0.000)	(0.000)	(0.000)
7. Craft, trade	-0.044***	-0.049***	-0.026***	0.033***
(Mean: 0.184)	(0.002)	(0.002)	(0.001)	(0.001)
8. Machine operators	-0.035***	-0.028***	-0.019***	0.021***
(Mean: 0.163)	(0.002)	(0.002)	(0.001)	(0.001)
9. Elementary	-0.009***	-0.010***	-0.007***	0.001
(Mean: 0.056)	(0.001)	(0.001)	(0.001)	(0.001)

Notes: Job upon labor market entry considered. Sample limited to Defence Forces sample due to data limitations. Shown at one-digit occupational level. Controls include demographic controls (age, year, native born, and language). Robust standard error are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Predictive power of skill measures: occupational choices (two-digit level)

	x_q	x_v	x_s	x_m
21. Science, engineering	0.024***	0.018***	-0.008***	-0.001
(Mean: 0.075)	(0.001)	(0.001)	(0.001)	(0.001)
22. Health	0.006***	0.008***	0.000	-0.005***
(Mean: 0.017)	(0.001)	(0.001)	(0.000)	(0.000)
23. Teaching	0.012***	0.010***	0.006***	-0.014***
(Mean: 0.037)	(0.001)	(0.001)	(0.001)	(0.001)
24. Business	0.008***	0.005***	0.009***	-0.006***
(Mean: 0.030)	(0.001)	(0.001)	(0.001)	(0.001)
25. IT	0.003***	0.004***	-0.001**	-0.001***
(Mean: 0.009)	(0.000)	(0.000)	(0.000)	(0.000)
26. Legal, social	0.002***	0.007***	0.004***	-0.010***
(Mean: 0.015)	(0.001)	(0.001)	(0.000)	(0.001)

Notes: Only professionals shown (one-digit code two). Job upon labor market entry considered. Sample limited to Defence Forces sample due to data limitations. Shown at two-digit occupational level. Controls include demographic controls (age, year, native born, and language). Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Predictive power of skill measures: wages

	Log income									
	All						BC		WC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
x_q	0.083***	0.032***	0.035***	0.034***			0.018***		0.041***	
	(0.002)	(0.002)	(0.002)	(0.002)			(0.003)		(0.005)	
x_v	0.014***	-0.004**	-0.003*	-0.002			-0.011***		0.010**	
	(0.002)	(0.002)	(0.002)	(0.002)			(0.004)		(0.005)	
x_s	0.059***	0.036***	0.039***	0.039***			0.029***		0.039***	
	(0.002)	(0.001)	(0.001)	(0.001)			(0.002)		(0.002)	
x_m	0.049***	0.022***	0.025***	0.024***			0.026***		0.018***	
	(0.002)	(0.001)	(0.001)	(0.001)			(0.002)		(0.002)	
y_q		0.468***	0.466***		0.192***		0.540***		0.418***	
		(0.011)	(0.011)		(0.012)		(0.016)		(0.020)	
y_v		-0.260***	-0.264***		-0.111***		-0.387***		-0.196***	
		(0.012)	(0.012)		(0.013)		(0.018)		(0.023)	
y_s		0.035***	0.037***		0.035***		0.027***		0.038***	
		(0.002)	(0.002)		(0.002)		(0.004)		(0.002)	
y_m		0.111***	0.117***		0.030***		0.164***		0.096***	
		(0.002)	(0.002)		(0.002)		(0.003)		(0.004)	
$x_q \cdot y_q$			0.031***	0.029***	0.029***	0.028***	0.011***	0.012**	0.024***	0.018***
			(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)	(0.005)	(0.004)
$x_v \cdot y_v$			0.026***	0.015***	0.021***	0.016***	0.011***	0.013**	0.009*	0.004
			(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)	(0.004)	(0.004)
$x_s \cdot y_s$			0.024***	0.021***	0.011***	0.009***	0.010***	0.003	0.027***	0.008***
			(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
$x_m \cdot y_m$			0.003***	-0.003**	0.004***	0.003**	-0.002	0.001	-0.007***	0.003*
			(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Observations	1,911,345	1,577,419	1,577,419	1,577,419	1,575,432	1,575,432	836,321	824,354	741,098	733,672
R-squared	0.147	0.188	0.193	0.255	0.623	0.629	0.085	0.642	0.222	0.651
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Occupation FE				Y		Y		Y		Y
Worker FE					Y	Y		Y		Y

Notes: Sample limited to Defence Forces sample and ages 24 or older. All available panel data included. Occupational skill requirements taken at two-digit level. Controls include demographic controls (age, year, native born, and language). Robust standard error clustered at the individual level and are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Occupational persistence by occupation

Occupation	Likelihood ratio
Panel A. One-digit level	
6. Agricultural, related workers (.011)	2.28
9. Elementary occupations (.049)	0.91
8. Plant and machine operators (.087)	0.8
2. Professionals (.234)	0.76
7. Craft and related trades workers (.11)	0.61
1. Managers (.02)	0.56
4. Clerical support workers (.061)	0.33
5. Service and sales workers (.203)	0.18
3. Technicians, assoc. professionals (.226)	0.18
Panel B. Two-digit level	
62. Skilled forestry, etc. workers (.001)	10.69
14. Hospitality, retail, other managers (.003)	7.11
96. Other elementary workers (.002)	5.96
44. Other clerical support workers (.011)	3.71
82. Assemblers (.009)	2.84
11. Executives, officials (.001)	2.73
26. Legal, social prof. (.031)	2.57
54. Protective services workers (.014)	2.56
61. Skilled agricultural (.01)	2.51
25. IT prof. (.018)	2.45
75. Food processing, other crafts (.011)	2.42
22. Health prof. (.022)	2.25
73. Handicraft, printing workers (.003)	2.03
94. Food prep assistants (.006)	1.76
83. Mobile plant operators (.044)	1.68
71. Building, other trades (.04)	1.6
91. Cleaners, helpers (.018)	1.5
81. Stationary machine operators (.034)	1.48
23. Teaching prof. (.054)	1.41
74. Electrical trades (.019)	1.37
93. Laborers in mining, etc. (.022)	1.25
72. Metal, machinery (.038)	1.05
35. Info., communications tech. (.015)	0.99
21. Science, engineer prof. (.064)	0.87
42. Customer services clerks (.019)	0.73
12. Commercial managers (.006)	0.71
43. Material recording clerks (.013)	0.7
52. Sales workers (.08)	0.53
51. Personal service workers (.051)	0.5
24. Business prof. (.043)	0.46
33. Business assoc. prof. (.068)	0.45
53. Personal care workers (.059)	0.44
13. Production, service managers (.009)	0.4
31. Science, engineer assoc. prof. (.05)	0.37
34. Legal, social assoc. prof. (.033)	0.35
32. Health assoc. prof. (.059)	0.26
41. General and keyboard clerks (.018)	0.14

Notes: Table shows how persistent each occupation is, at each occupational classification. Likelihood ratio is equal to coefficient β_{1j} divided by the control mean in the given occupation. Sorted within each occupational classification by likelihood ratio. Sample frequency mean in parentheses.

Table A6: Occupational persistence by occupation (cont.)

Occupation	Likelihood ratio	Occupation	Likelihood ratio
Panel C. Three-digit level			
613. Mixed crop and animal producers (.001)	34.04	235. Other teaching prof. (.011)	2.4
516. Other personal services (.002)	22.67	252. Database and network prof. (.002)	2.39
141. Managers, Hotel & restaurant (.001)	21.66	711. Building frame workers (.024)	2.36
315. Ship and aircraft tech. (.002)	18.47	523. Cashiers and ticket clerks (.005)	2.3
261. Legal prof. (.006)	16.62	833. Heavy truck and bus drivers (.016)	2.26
751. Food processing workers (.002)	16.2	741. Electrical equipment installers, repairers (.014)	2.25
265. Artists (.006)	14.34	233. Secondary ed. teachers (.013)	2.22
612. Animal producers (.004)	13.42	933. Transport and storage labourers (.017)	2.21
142. Retail managers (.001)	12.11	722. Blacksmiths, toolmakers , etc. (.011)	2.16
511. Travel guides etc. (.002)	10.82	232. Vocational teachers (.003)	2.09
813. Chemical products plant, machine operators (.003)	10.5	515. Building and housekeeping supervisors (.009)	2.09
817. Wood processing & papermaking plant operators (.006)	10.41	234. Primary ed. & early childhood teachers (.023)	1.97
812. Metal processing & finishing plant operators (.003)	9.88	121. Business, service managers (.003)	1.96
611. Market gardeners and crop workers (.003)	9.41	333. Business services agents (.009)	1.8
713. Painters, building structure cleaners, etc. (.004)	9.35	941. Food preparation assistants (.006)	1.78
221. Medical doctors (.011)	9.22	242. Administration prof. (.012)	1.74
962. Other elementary workers (.002)	7.12	335. Regulatory gov assoc. prof. (.009)	1.69
811. Mining & mineral processing plant operators (.003)	7.06	723. Machinery mechanics and repairers (.019)	1.63
816. Food products machine operators (.007)	6.89	321. Medical tech. (.008)	1.6
753. Garment workers (.002)	6.0	513. Waiters & bartenders (.009)	1.58
831. Locomotive engine drivers (.001)	5.93	911. Hotel and office cleaners (.018)	1.49
815. Textile products machine operators (.003)	5.72	215. Electrotechnology engineers (.023)	1.37
514. Hairdressers, beauticians, etc. (.008)	5.39	432. Material-recording & transport clerks (.006)	1.36
752. Wood treaters, cabinet-makers, etc. (.005)	5.11	421. Tellers, clerks, etc. (.009)	1.33
932. Manufacturing labourers (.001)	4.55	331. Finance, math, assoc. prof. (.014)	1.31
712. Building finishers (.012)	4.47	332. Sales agents, brokers (.025)	1.27
262. Librarians (.001)	4.42	343. Artistic, cultural assoc. prof. (.006)	1.25
834. Mobile plant operators (.014)	4.34	134. Professional services managers (.004)	1.18
814. Rubber, plastic & paper products machine operators (.004)	4.15	243. Sales, marketing, PR prof. (.022)	1.03
832. Car, van and motorcycle drivers (.012)	3.97	214. Engineering prof. (.023)	0.96
264. Authors, etc. (.007)	3.88	314. Life science tech, etc. (.003)	0.94
133. IT, service managers (.001)	3.88	351. IT tech. (.013)	0.92
818. Other stationary plant & machine operators (.006)	3.74	325. Other health assoc. prof. (.015)	0.87
132. Managers, manufacturing (.004)	3.63	312. Supervisors, manufacturing, construction, etc. (.005)	0.81
441. Other clerical support (.011)	3.52	431. Numerical clerks (.008)	0.77
216. Architects, designers, etc. (.009)	3.5	522. Shop salespersons (.056)	0.75
931. Mining and construction labourers (.004)	3.44	531. Child care & teachers' aides (.022)	0.71
211. Physical & earth science prof. (.003)	3.27	341. Legal, social assoc. prof. (.021)	0.69
213. Life science prof. (.005)	3.17	422. Client information workers (.01)	0.68
352. Telecommunication tech. (.002)	3.17	311. Physical and engineering tech. (.038)	0.58
754. Other craft & related workers (.002)	3.13	742. Electronics and telecomm installers, repairers (.005)	0.58
821. Assemblers (.009)	3.04	532. Personal care in health services (.038)	0.55
263. Social and religious prof. (.012)	2.98	512. Cooks (.018)	0.55
231. Higher ed. teachers (.004)	2.87	521. Street and market salespersons (.008)	0.5
342. Sports and fitness (.006)	2.77	524. Other sales workers (.013)	0.47
251. Software developers (.015)	2.67	122. Sales & marketing managers (.003)	0.47
721. Sheet metal workers, moulders, welders, etc. (.008)	2.65	334. Administrative secretaries (.012)	0.39
732. Printing trades workers (.003)	2.63	322. Nursing assoc. prof. (.037)	0.3
226. Other health prof. (.01)	2.59	411. General office clerks (.007)	0.25
313. Process control tech. (.003)	2.51	241. Finance prof. (.01)	0.18
111. Legislators, officials (.001)	2.48	412. Secretaries (.01)	0.17
541. Protective service workers (.014)	2.47		

Notes: See table notes from Panels A and B.

Table A6: Occupational persistence by occupation (cont.)

Occupation	Likelihood ratio	Occupation	Likelihood ratio
Panel D. Four-digit level			
8181. Glass and ceramics plant operators (.001)	53.15	2619. Other legal prof. (.002)	8.69
7511. Butchers, fishmongers and related food preparers (.001)	47.33	3339. Other business services agents (.003)	8.46
6111. Field crop and vegetable growers (.001)	34.62	4412. Mail carriers and sorting clerks (.007)	8.38
6130. Mixed crop and animal producers (.001)	33.31	8219. Other assemblers (.002)	8.27
2262. Pharmacists (.001)	32.82	9329. Other manufacturing labourers (.002)	8.15
8343. Crane, hoist and related plant operators (.001)	30.79	1321. Manufacturing managers (.001)	8.05
2250. Veterinarians (.001)	29.34	5412. Police officers (.004)	8.02
8341. Mobile farm and forestry plant operators (.003)	28.24	3352. Government tax and excise officials (.001)	7.75
2652. Musicians, singers and composers (.002)	27.67	3132. Incinerator and water treatment plant operators (.002)	7.7
7124. Insulation workers (.002)	27.45	5221. Shop keepers (.002)	7.64
2261. Dentists (.002)	25.96	2432. Public relations prof. (.002)	7.48
7512. Bakers, pastry-cooks and confectionery makers (.001)	25.45	5141. Hairdressers (.006)	7.33
8114. Mineral products machine operators (.002)	25.29	2352. Special needs teachers (.003)	7.17
8121. Metal processing plant operators (.002)	24.09	3211. Medical imaging and therapeutic equipment tech. (.002)	7.1
8172. Wood processing plant operators (.002)	22.23	9312. Civil engineering labourers (.002)	6.92
2166. Graphic and multimedia designers (.003)	20.58	7126. Plumbers and pipe fitters (.007)	6.89
8312. Railway brake, signal and switch operators (.001)	20.29	8160. Food and related products machine operators (.007)	6.83
2611. Lawyers (.004)	19.41	3355. Police inspectors and detectives (.001)	6.64
2654. Film, stage and related directors and producers (.001)	18.79	2642. Journalists (.004)	6.57
2636. Religious professionals (.001)	18.35	7411. Building and related electricians (.007)	6.33
3521. Broadcasting and AV tech. (.001)	17.62	7119. Other building frame and related trades workers (.002)	6.31
2111. Physicists and astronomers (.001)	17.56	7522. Cabinet-makers and related workers (.002)	6.19
6121. Livestock and dairy producers (.004)	16.81	8183. Packing, bottling and labelling machine operators (.003)	6.13
8171. Pulp and papermaking plant operators (.004)	16.68	8212. Electrical and electronic equipment assemblers (.004)	6.07
2212. Specialist medical practitioners (.006)	14.56	2132. Farming, forestry and fisheries advisers (.002)	5.94
7131. Painters and related workers (.003)	14.01	2161. Building architects (.003)	5.84
1323. Construction managers (.002)	14.0	2145. Chemical engineers (.002)	5.82
3431. Photographers (.002)	13.79	8211. Mechanical machinery assemblers (.004)	5.63
9621. Messengers, package deliverers and luggage porters (.001)	13.65	2643. Translators, interpreters and other linguists (.002)	5.35
7322. Printers (.001)	12.92	3321. Insurance representatives (.002)	5.32
3331. Clearing and forwarding agents (.001)	12.7	3213. Pharmaceutical tech. and assistants (.003)	5.22
3351. Customs and border inspectors (.003)	12.32	1330. IT service managers (.001)	5.03
5411. Fire-fighters (.003)	11.83	3423. Fitness and recreation instructors and program leaders (.003)	4.95
8142. Plastic products machine operators (.003)	11.8	2133. Environmental protection professionals (.001)	4.93
5151. Cleaning and housekeeping supervisors (.001)	11.71	8332. Heavy truck and lorry drivers (.015)	4.88
7413. Electrical line installers and repairers (.001)	11.69	8344. Lifting truck operators (.004)	4.78
1420. Retail and wholesale trade managers (.001)	11.6	4221. Travel consultants and clerks (.002)	4.73
8131. Chemical products plant and machine operators (.003)	11.44	5164. Pet groomers and animal care workers (.001)	4.71
8322. Car, taxi and van drivers (.003)	11.09	3353. Government social benefits officials (.003)	4.7
2632. Sociologists, anthropologists, etc. (.001)	10.68	7212. Welders and flamecutters (.003)	4.68
2634. Psychologists (.003)	10.63	4312. Statistical, finance and insurance clerks (.003)	4.52
8342. Earthmoving and related plant operators (.006)	10.59	2151. Electrical engineers (.006)	4.42
6113. Gardeners, horticultural and nursery growers (.002)	10.54	7213. Sheet-metal workers (.003)	4.39
5245. Service station attendants (.001)	10.43	3259. Other health asso prof. (.001)	4.39
8331. Bus and tram drivers (.002)	10.37	2163. Product and garment designers (.002)	4.25
7523. Woodworking-machine tool setters and operators (.002)	10.01	5414. Security guards (.006)	4.25
5142. Beauticians and related workers (.002)	9.85	7422. IT installers and servicers (.003)	4.08
2211. Generalist medical practitioners (.005)	9.64	4226. Receptionists (general) (.001)	3.98
3413. Religious assoc. prof. (.002)	9.52	7223. Metal working machine tool setters and operators (.007)	3.94
2120. Mathematicians, actuaries and statisticians (.001)	9.49	7115. Carpenters and joiners (.005)	3.73
3334. Real estate agents and property managers (.003)	9.36	2412. Financial and investment advisers (.002)	3.72
2514. Applications programmers (.001)	9.17		

Notes: See table notes from Panels A and B.

Table A6: Occupational persistence by occupation (cont.)

Occupation	Likelihood ratio	Occupation	Likelihood ratio
Panel D. Four-digit level (cont.)			
2131. Biologists, botanists, zoologists, etc. (.002)	3.68	3123. Construction supervisors (.003)	1.66
2421. Management and organization analysts (.003)	3.63	3421. Athletes and sports players (.002)	1.65
2423. Personnel and careers prof. (.003)	3.56	3122. Manufacturing supervisors (.001)	1.63
2341. Primary school teachers (.015)	3.46	4313. Payroll clerks (.002)	1.62
7231. Motor vehicle mechanics and repairers (.01)	3.42	9112. Cleaners and helpers (.018)	1.59
7543. Product graders and testers (excluding foods and beverages) (.002)	3.17	3312. Credit and loans officers (.002)	1.57
3116. Chemical engineering tech. (.002)	3.12	2152. Electronics engineers (.005)	1.57
1219. Other business services and administration managers (.002)	3.07	2342. Early childhood educators (.01)	1.53
2113. Chemists (.001)	3.05	3114. Electronics engineering tech. (.006)	1.51
2434. IT sales prof. (.002)	2.97	3212. Medical and pathology laboratory tech. (.002)	1.48
5246. Food service counter attendants (.004)	2.9	2431. Advertising and marketing prof. (.011)	1.48
2310. University and higher education teachers (.004)	2.78	2351. Education methods specialists (.002)	1.46
3333. Employment agents and contractors (.001)	2.75	7421. Electronics mechanics and servicers (.003)	1.45
7111. House builders (.01)	2.75	5329. Other personal care workers in health services (.004)	1.38
2330. Secondary education teachers (.013)	2.73	2149. Other engineering prof. (.003)	1.37
5131. Waiters (.008)	2.73	2422. Policy administration prof. (.005)	1.37
5222. Shop supervisors (.006)	2.7	3115. Mechanical engineering tech. (.011)	1.35
2411. Accountants (.008)	2.65	5244. Contact centre salespersons (.003)	1.31
1221. Sales and marketing managers (.002)	2.65	2141. Industrial and production engineers (.003)	1.18
7222. Toolmakers and related workers (.002)	2.63	3258. Ambulance workers (.005)	1.13
9313. Building construction labourers (.001)	2.56	5312. Teachers' aides (.006)	1.11
4411. Library clerks (.001)	2.53	2519. Other software and applications developers & analysts (.002)	1.1
3113. Electrical engineering tech. (.003)	2.48	3341. Office supervisors (.001)	0.99
4323. Transport clerks (.004)	2.44	5322. Home-based personal care workers (.008)	0.97
5230. Cashiers and ticket clerks (.005)	2.42	5311. Child care workers (.017)	0.96
3332. Conference and event planners (.001)	2.4	3412. Social work assoc. prof. (.018)	0.94
2635. Social work and counselling prof. (.005)	2.36	4321. Stock clerks (.002)	0.93
1346. Financial and insurance services branch managers (.003)	2.31	3511. IT operations tech. (.01)	0.9
3323. Buyers (.004)	2.3	5120. Cooks (.018)	0.87
2511. Systems analysts (.004)	2.27	2433. Technical and medical sales prof. (.008)	0.85
5153. Building caretakers (.008)	2.25	2266. Audiologists and speech therapists (.002)	0.8
3434. Chefs (.002)	2.24	3222. Midwifery assoc. prof. (.002)	0.8
7233. Agricultural and industrial machinery mechanics and repairers (.007)	2.23	5211. Stall and market salespersons (.007)	0.76
9333. Freight handlers (.018)	2.19	5321. Health care assistants (.028)	0.71
7412. Electrical mechanics and fitters (.006)	2.19	5223. Shop sales assistants (.049)	0.65
2512. Software developers (.01)	2.17	2164. Town and traffic planners (.001)	0.6
2359. Other teaching prof. (.004)	2.15	3221. Nursing assoc. prof. (.037)	0.57
2320. Vocational education teachers (.003)	2.07	4311. Accounting and bookkeeping clerks (.004)	0.55
2144. Mechanical engineers (.01)	2.05	4229. Other client information workers (.002)	0.48
3141. Life science technicians (.002)	2.02	3119. Other physical and engineering science tech. (.009)	0.42
3411. Legal and Related assoc. prof. (.002)	1.98	3343. Administrative and executive secretaries (.004)	0.37
3313. Accounting assoc. prof. (.01)	1.98	3512. IT user supporttech. (.006)	0.29
4224. Hotel receptionists (.001)	1.98	4110. General office clerks (.007)	0.21
3255. Physiotherapy tech. and assistants (.006)	1.91	4120. Secretaries (general) (.012)	0.19
9412. Kitchen helpers (.005)	1.85	3118. Draughtspersons (.001)	0.01
2263. Environmental and occupational health and hygiene prof. (.005)	1.83	4222. Contact centre information clerks (.002)	-0.23
3344. Medical secretaries (.002)	1.82	2153. Telecommunications engineers (.001)	-0.98
3322. Commercial sales representatives (.019)	1.74	3432. Interior designers and decorators (.001)	-1.02
3112. Civil engineering tech. (.007)	1.72	3251. Veterinary tech. and assistants (.001)	-1.04
4211. Bank tellers and related clerks (.008)	1.71	4225. Enquiry clerks (.001)	-1.14
2142. Civil engineers (.005)	1.7		
4419. Other clerical support workers (.003)	1.68		

Notes: See table notes from Panels A and B.

Table A7: Intergenerational occupational persistence: by parent and by gender

	Choose occupation j								
	Both					Sons		Daughters	
	Primary (1)	Union (2)	Same-sex (3)	Father (4)	Mother (5)	Father (6)	Mother (7)	Father (8)	Mother (9)
Panel A. Class level									
Parent in occupation j	0.145	0.143	0.135	0.133	0.131	0.138	0.129	0.127	0.132
	(0.0019)	(0.0017)	(0.0020)	(0.0020)	(0.0020)	(0.0028)	(0.0028)	(0.0029)	(0.0029)
Observations (millions)	0.5	0.5	0.5	0.4	0.5	0.2	0.2	0.2	0.2
Mean	0.4105	0.4107	0.4110	0.4107	0.4121	0.4006	0.3793	0.4206	0.4054
Percent increase	35%	35%	33%	32%	32%	35%	34%	30%	33%
Panel B. One-digit level									
Parent in occupation j	0.0584	0.0538	0.0675	0.0545	0.0542	0.0702	0.0437	0.0387	0.0641
	(0.0009)	(0.0007)	(0.0010)	(0.0010)	(0.0010)	(0.0014)	(0.0013)	(0.0012)	(0.0014)
Observations (millions)	2.3	2.3	2.1	2.0	2.1	1.0	1.1	1.0	1.1
Mean	0.1015	0.1015	0.1013	0.1015	0.1014	0.0989	0.1031	0.1040	0.0946
Percent increase	58%	53%	67%	54%	53%	71%	42%	37%	68%
Panel C. Two-digit level									
Parent in occupation j	0.0380	0.0346	0.0496	0.0358	0.0353	0.0533	0.0250	0.0185	0.0452
	(0.0005)	(0.0004)	(0.0006)	(0.0006)	(0.0006)	(0.0009)	(0.0007)	(0.0007)	(0.0009)
Observations (millions)	9.8	10.0	8.9	8.5	9.3	4.2	4.6	4.3	4.7
Mean	0.0244	0.0244	0.0244	0.0244	0.0244	0.0238	0.0249	0.0251	0.0236
Percent increase	155%	142%	203%	146%	144%	224%	100%	74%	192%
Panel D. Three-digit level									
Parent in occupation j	0.0292	0.0264	0.0386	0.0286	0.0262	0.0444	0.0194	0.0129	0.0327
	(0.0009)	(0.0007)	(0.0011)	(0.0010)	(0.0010)	(0.0017)	(0.0011)	(0.0010)	(0.0016)
Observations (millions)	5.9	6.0	5.3	5.0	5.6	2.5	2.7	2.5	2.8
Mean	0.0081	0.0081	0.0081	0.0081	0.0081	0.0079	0.0082	0.0082	0.0079
Percent increase	363%	328%	479%	354%	325%	562%	238%	156%	415%
Panel E. Four-digit level									
Parent in occupation j	0.0266	0.0237	0.0346	0.0241	0.0251	0.0387	0.0189	0.0103	0.0308
	(0.0009)	(0.0006)	(0.0010)	(0.0009)	(0.0009)	(0.0016)	(0.0011)	(0.0009)	(0.0015)
Observations (millions)	18.1	18.6	16.1	14.7	17.2	7.1	8.3	7.5	8.9
Mean	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0024
Percent increase	1065%	951%	1388%	968%	1005%	1577%	754%	406%	1266%

Notes: Job upon labor market entry of all children shown. Demographic controls include sex, age, year of birth, native born, and language. Class level, one-digit, and two-digit specifications limited to 1975-1979 cohorts for computational feasibility. Three-digit and four-digit specifications limited to 1978 cohort only for computational feasibility. Standard errors in parentheses.

Table A8: Other heterogeneity: predictors of occupational following

	Occupational follower			
	All	Defence Forces		
		All	Blue collar	White collar
	(1)	(2)	(3)	(4)
Demographics				
Male	-0.0061***			
	(0.0003)			
Age	-0.0003***	0.0009***	0.0001	0.0027***
	(0.0001)	(0.0002)	(0.0003)	(0.0003)
Urban	-0.0170***	-0.0239***	-0.0331***	-0.0013
	(0.0005)	(0.0007)	(0.0009)	(0.0010)
Native born	0.0072***	0.0128***	0.0125***	0.0059
	(0.0011)	(0.0021)	(0.0026)	(0.0036)
Finnish language	-0.0072***	0.0013	0.0052***	-0.0035*
	(0.0007)	(0.0014)	(0.0020)	(0.0018)
Adopted	-0.0584***	-0.0502***	-0.0394***	-0.0674***
	(0.0020)	(0.0058)	(0.0074)	(0.0039)
Parent characteristics				
Parental income	0.0078***			
	(0.0004)			
Father income		0.0017***	0.0100***	-0.0004***
		(0.0003)	(0.0005)	(0.0001)
Skills, skill mismatch, and education				
x_q		-0.0034***	-0.0076***	0.0040***
		(0.0004)	(0.0005)	(0.0005)
x_v		-0.0062***	-0.0108***	0.0020***
		(0.0004)	(0.0005)	(0.0006)
x_s		-0.0077***	-0.0082***	-0.0008*
		(0.0003)	(0.0004)	(0.0004)
x_m		0.0032***	0.0059***	0.0034***
		(0.0003)	(0.0004)	(0.0004)
Horizontal mismatch		-0.0168***	-0.0095***	-0.0077***
		(0.0003)	(0.0005)	(0.0005)
Vertical mismatch		-0.0011***	0.0023***	-0.0009**
		(0.0003)	(0.0004)	(0.0005)
College	-0.0369***	-0.0310***	-0.1037***	0.0662***
	(0.0004)	(0.0006)	(0.0007)	(0.0008)
Observations	4,354,269	1,332,967	761,573	571,394
Parent	Either	Father	Father	Father
Follower rate	0.1267	0.1076	0.1182	0.0934

Notes: Sample starts at age 21, and all years after included. Limited to cohorts born between 1975-1980 for computational tractability. Year fixed effect included. Shown at 2-digit occupational level. Both sons and daughters considered in overall sample, with respect to union of parents. Only fathers considered in Defence Forces sample. Age in 10's, parent income in 10,000's. Skill and mismatch measures standardized. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A9: Intergenerational occupational persistence: statistical decomposition

	Choose occupation j											
	All Occupations				WC Occupations				BC Occupations			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A. Class level												
Parent in occupation j	0.159	0.136	0.100	0.0647								
	(0.0030)	(0.0030)	(0.0030)	(0.0028)								
Mean	0.3906	0.3906	0.3906	0.3906								
Percent increase	41%	35%	26%	17%								
Persistence explained		14%	37%	59%								
Panel B. One-digit level												
Parent in occupation j	0.0763	0.0712	0.0623	0.0511	0.0581	0.0517	0.0395	0.0229	0.0909	0.0868	0.0803	0.0734
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0021)	(0.0021)	(0.0021)	(0.0020)
Mean	0.0975	0.0975	0.0975	0.0975	0.1064	0.1064	0.1064	0.1064	0.0933	0.0933	0.0933	0.0933
Percent increase	78%	73%	64%	52%	55%	49%	37%	22%	98%	93%	86%	79%
Persistence explained		7%	18%	33%		11%	32%	61%		5%	12%	19%
Panel C. Two-digit level												
Parent in occupation j	0.0584	0.0569	0.0543	0.0472	0.0302	0.0286	0.0258	0.0141	0.0812	0.0797	0.0773	0.0739
	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Mean	0.0236	0.0236	0.0236	0.0236	0.0225	0.0225	0.0225	0.0225	0.0242	0.0242	0.0242	0.0242
Percent increase	248%	241%	230%	200%	134%	127%	115%	63%	335%	329%	319%	305%
Persistence explained		3%	7%	19%		5%	15%	53%		2%	5%	9%
Panel D. Three-digit level												
Parent in occupation j	0.0492	0.0487	0.0478	0.0431	0.0196	0.0191	0.0183	0.0100	0.0755	0.0750	0.0739	0.0723
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0021)	(0.0021)	(0.0021)	(0.0019)	(0.0030)	(0.0030)	(0.0030)	(0.0030)
Mean	0.0079	0.0079	0.0079	0.0079	0.0062	0.0062	0.0062	0.0062	0.0093	0.0093	0.0093	0.0093
Percent increase	624%	617%	605%	545%	316%	307%	295%	162%	808%	803%	791%	773%
Persistence explained		1%	3%	13%		3%	7%	49%		1%	2%	4%
Panel E. Four-digit level												
Parent in occupation j	0.0389	0.0387	0.0383	0.0354	0.0169	0.0167	0.0163	0.0109	0.0591	0.0589	0.0584	0.0578
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0016)	(0.0028)	(0.0028)	(0.0028)	(0.0028)
Mean	0.0026	0.0026	0.0026	0.0026	0.0019	0.0019	0.0019	0.0019	0.0032	0.0032	0.0032	0.0032
Percent increase	1525%	1517%	1501%	1388%	875%	866%	848%	567%	1864%	1857%	1842%	1823%
Persistence explained		1%	2%	9%		1%	3%	35%		0%	1%	2%
Controls	Raw	1-D Skill	M-D Skills	Major	Raw	1-D Skill	M-D Skills	Major	Raw	1-D Skill	M-D Skills	Major

Notes: Job upon labor market entry of sons considered with respect to fathers, due to fact that skill data is limited to males. Demographic controls include age, year of birth, native born, and language. Class level, one-digit, and two-digit specifications limited to 1975-1979 cohorts for computational feasibility. Three-digit and four-digit specifications limited to 1978 cohort only for computational feasibility. Standard errors in parentheses.

Table A10: Intergenerational occupational persistence: by one-digit occupation

	Choose occupation j			
	(1)	(2)	(3)	(4)
1. Managers *(Mean: 0.0116)*				
Parent in occupation j	0.0136*** (0.0016)	0.0128*** (0.0016)	0.0121*** (0.0016)	0.0106*** (0.0016)
Percent increase	117%	110%	104%	91%
Persistence explained		6%	11%	23%
2. Professionals *(Mean: 0.1381)*				
Parent in occupation j	0.114*** (0.0037)	0.103*** (0.0037)	0.0745*** (0.0037)	0.0389*** (0.0035)
Percent increase	82%	74%	54%	28%
Persistence explained		10%	34%	66%
3. Assoc. professionals *(Mean: 0.1767)*				
Parent in occupation j	0.0339*** (0.0035)	0.0288*** (0.0035)	0.0250*** (0.0035)	0.0162*** (0.0034)
Percent increase	19%	16%	14%	9%
Persistence explained		15%	26%	52%
4. Clerical support workers *(Mean: 0.0417)*				
Parent in occupation j	0.0470*** (0.0058)	0.0471*** (0.0058)	0.0464*** (0.0058)	0.0462*** (0.0058)
Percent increase	113%	113%	111%	110%
Persistence explained		0%	1%	2%
5. Sales workers *(Mean: 0.1313)*				
Parent in occupation j	0.0367*** (0.0034)	0.0367*** (0.0034)	0.0368*** (0.0034)	0.0331*** (0.0034)
Percent increase	28%	28%	28%	25%
Persistence explained		0%	0%	10%
6. Agricultural *(Mean: 0.0056)*				
Parent in occupation j	0.0194*** (0.0044)	0.0188*** (0.0044)	0.0174*** (0.0044)	0.0167*** (0.0044)
Percent increase	346%	335%	310%	298%
Persistence explained		3%	10%	14%
7. Craft workers *(Mean: 0.1650)*				
Parent in occupation j	0.116*** (0.0035)	0.108*** (0.0035)	0.0951*** (0.0034)	0.0846*** (0.0034)
Percent increase	70%	66%	58%	51%
Persistence explained		7%	18%	27%
8. Machine operators *(Mean: 0.1608)*				
Parent in occupation j	0.143*** (0.0043)	0.137*** (0.0043)	0.130*** (0.0043)	0.121*** (0.0043)
Percent increase	89%	85%	81%	75%
Persistence explained		4%	9%	15%
9. Elementary occupations *(Mean: 0.0820)*				
Parent in occupation j	0.0463*** (0.0067)	0.0450*** (0.0067)	0.0442*** (0.0067)	0.0429*** (0.0067)
Percent increase	56%	55%	54%	52%
Persistence explained		3%	4%	7%
Observations	93,107	93,107	93,107	93,107
Controls	Raw	1-D Skill	M-D Skills	Major

Notes: Job upon labor market entry of sons considered with respect to fathers, due to fact that skill data is limited to males. Demographic controls include age, year of birth, native born, and language. Limited to 1975-1979 cohorts for computational feasibility. Standard errors in parentheses.

Table A11: Intergenerational occupational persistence: by two-digit occupation (professional occupations only)

	Choose occupation j			
	(1)	(2)	(3)	(4)
21. Science and engineering professionals *(Mean: 0.0691)*				
Parent in occupation j	0.0422*** (0.0044)	0.0376*** (0.0045)	0.0297*** (0.0045)	0.0136*** (0.0043)
Percent increase	61%	54%	43%	20%
Persistence explained		11%	30%	68%
22. Health professionals *(Mean: 0.0120)*				
Parent in occupation j	0.0623*** (0.0070)	0.0604*** (0.0070)	0.0560*** (0.0070)	0.0138*** (0.0047)
Percent increase	520%	504%	467%	116%
Persistence explained		3%	10%	78%
23. Teaching professionals *(Mean: 0.0277)*				
Parent in occupation j	0.0607*** (0.0049)	0.0585*** (0.0049)	0.0522*** (0.0048)	0.0255*** (0.0041)
Percent increase	220%	211%	188%	92%
Persistence explained		4%	14%	58%
24. Business professionals *(Mean: 0.0261)*				
Parent in occupation j	0.0211*** (0.0037)	0.0198*** (0.0037)	0.0173*** (0.0037)	0.0120*** (0.0037)
Percent increase	81%	76%	66%	46%
Persistence explained		6%	18%	43%
25. IT professionals *(Mean: 0.0064)*				
Parent in occupation j	0.0148* (0.0077)	0.0145* (0.0077)	0.0131* (0.0077)	0.0104 (0.0078)
Percent increase	231%	226%	204%	162%
Persistence explained		2%	12%	30%
26. Legal, social professionals *(Mean: 0.0121)*				
Parent in occupation j	0.0518*** (0.0067)	0.0515*** (0.0067)	0.0455*** (0.0067)	0.0183*** (0.0056)
Percent increase	426%	424%	374%	151%
Persistence explained		0%	12%	65%
Observations	93,107	93,107	93,107	93,107
Controls	Raw	1-D Skill	M-D Skills	Major

Notes: Shown for all two-digit occupations that are in the professional occupations (one-digit code 2). Job upon labor market entry of sons considered with respect to fathers, due to fact that skill data is limited to males. Demographic controls include age, year of birth, native born, and language. Limited to 1975-1979 cohorts for computational feasibility. Standard errors in parentheses.

Table A12: Amount of wage variation explained by occupational choice

Occupational classification	Share explained	R^2	Increase
None		0.093	
Class	0.102	0.213	0.120
1-Digit	0.139	0.254	0.161
2-Digit	0.193	0.310	0.217
3-Digit	0.233	0.350	0.257
4-Digit	0.265	0.382	0.289

Notes: Taken from cross-sectional regression at age 35, with gender and year controls. Share explained is proportion of wage variation explained by occupation fixed effects. R^2 is amount of variation explained by model. Overall increase is increase in R^2 compared to case without occupation fixed effects.

Table A13: Intergenerational mobility and occupational choice: by occupational classification

	Child income rank				
	(1)	(2)	(3)	(4)	(5)
Parent income rank	0.156***	0.097***	0.079***	0.072***	0.064***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	42.211***	45.154***	46.036***	46.396***	46.812***
	(0.088)	(0.083)	(0.077)	(0.074)	(0.072)
Observations	424,744	424,744	424,744	424,744	424,744
Occupation FE		Y	Y	Y	Y
Digit-level		1	2	3	4
Share explained		37.8%	49.1%	53.7%	59.1%

Notes: All children included, with respect to total parental income. Using mean child income from ages 25-35, and then ranked within birth cohort. Only children for which I observe until age 35 included, and so therefore limited to cohorts born between 1975 and 1983. Current occupation at age 35 used. Total parental income taken as average income over ages 25-55, and then ranked within child birth cohort. Occupational classification indicated in table. Share explained is share of regression coefficient in Column (1) explained by occupation fixed effect in respective column. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A14: The effect of occupational following on future income

	Income			
	Income	Income 1 year later	Income 2 years later	Income 3 years later
	(1)	(2)	(3)	(4)
Panel A. IV estimates				
Same occupation	1,928.620**	2,694.766*	3,443.355*	1,414.718
	(822.717)	(1,418.282)	(1,882.623)	(2,042.189)
Panel B. OLS estimates				
Same occupation	447.660***	879.792***	1,000.903***	998.857***
	(26.270)	(46.332)	(56.809)	(64.125)
First stage F-statistic	677	659	628	594
Outcome mean	2.1e+04	2.4e+04	2.5e+04	2.6e+04
Observations	621,444	606,748	579,148	558,408

Notes: Main specification (Column (3) of Table 4) shown. All children included with respect to primary earning parent. Occupation taken at labor market entry. Two-digit occupation shown. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A15: The effect of occupational following on income: by occupational class

	Log income					
	All		BC		WC	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Same occupation	0.018***	0.055*	-0.015***	0.013	0.058***	0.095
	(0.001)	(0.031)	(0.001)	(0.030)	(0.002)	(0.076)
Partial F-statistic		677		512		155
Following rate	0.083	0.083	0.095	0.095	0.070	0.070
Observations	621,444	621,444	334,332	334,332	287,110	287,110

Notes: Main specification (Column (3) of Table 4) shown. Sample split by whether parent is from a blue collar or white collar occupation. All children included with respect to primary earning parent. Occupation taken at labor market entry. Two-digit occupation shown. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A16: The effect of occupational following on income: by sex

	Log income			
	Father		Mother	
	Sons (1)	Daughters (2)	Sons (3)	Daughter (4)
Same occupation	0.032	0.030	0.069	0.061
	(0.058)	(0.089)	(0.089)	(0.045)
First stage partial F-statistic	161	117	155	211
Following rate	0.112	0.051	0.049	0.114
Observations	170,026	156,013	135,450	159,942

Notes: Main specification (Column (3) of Table 4) shown. Sample split by sex of parent and sex of child. Occupation taken at labor market entry. Two-digit occupation shown. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A17: The effect of occupational following on income: by occupational classification

	Log income			
	(1)	(2)	(3)	(4)
Same occupation	-0.054*	0.055*	0.141***	0.196***
	(0.030)	(0.031)	(0.041)	(0.044)
Occupational classification (digit-level)	1	2	3	4
First stage partial F-statistic	361	677	635	730
Following rate	0.214	0.083	0.047	0.035
Observations	622,661	621,444	611,348	566,129

Notes: Main specification (Column (3) of Table 4) shown. Shown at each occupational classification separately. Occupation taken at labor market entry. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A18: The effect of occupational following on income: placebo check

	First stage		Reduced form		2SLS	
	Actual	Placebo	Actual	Placebo	Actual	Placebo
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Two-digit level						
Hiring rate	0.108***	0.004	0.006*	-0.003		
	(0.005)	(0.004)	(0.003)	(0.003)		
Same occupation					0.054*	-0.856
					(0.032)	(1.234)
First stage partial F-statistic	644	1				
Hiring rate mean	0.097	0.088				
Following rate	0.083	0.083				
Observations	608,211	606,163	608,211	606,163	608,211	606,163
Panel B. Three-digit level						
Hiring rate	0.140***	-0.002	0.020***	0.010		
	(0.007)	(0.007)	(0.006)	(0.006)		
Same occupation					0.142***	-5.105
					(0.041)	(16.634)
First stage partial F-statistic	628	0				
Hiring rate mean	0.040	0.029				
Following rate	0.047	0.046				
Observations	602,630	572,142	602,630	572,142	602,630	572,142
Panel C. Four-digit level						
Hiring rate	0.145***	0.052***	0.028***	0.000		
	(0.008)	(0.010)	(0.006)	(0.016)		
Same occupation					0.196***	0.004
					(0.044)	(0.301)
First stage partial F-statistic	730	17				
Hiring rate mean	0.025	0.011				
Following rate	0.035	0.035				
Observations	566,129	477,037	566,129	477,037	566,129	477,037

Notes: Main specification (Column (3) of Table 4) shown. Shown at each occupational classification separately. Occupation taken at labor market entry. Demographic controls include age, sex, native born, language, and sex and income of primary earner. "Actual" columns refer to actual results. "Placebo" columns refer to placebo results, where hiring rate is replaced by randomly-chosen hiring rate from a different occupation in the same coarser occupation (e.g., for the three-digit classification, another three-digit occupation from the same two-digit occupation is chosen). Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A19: The effect of occupational following on income: controlling for same firm

	Log income				
	(1)	(2)	(3)	(4)	(5)
Same occupation	0.018***	0.020***	0.055*	0.065**	
	(0.001)	(0.001)	(0.031)	(0.033)	
Same firm		0.015***		-0.010	
		(0.002)		(0.010)	
Same occupation, different firm					0.079**
					(0.038)
Specification	OLS	OLS	IV	IV	IV
Partial F-statistic			677	652	561
Occ. following rate	0.083	0.083	0.083	0.083	0.067
Firm following rate		0.045		0.046	0.046
Observations	621,444	548,242	621,444	603,702	603,702

Notes: Main specification (Column (3) of Table 4) shown. Occupation taken at labor market entry. Demographic controls include age, sex, native born, language, and sex and income of primary earner. Columns (2) and (4) include a control for the same firm as parent. Column (5) defines the endogenous variable as same occupation but different from from parent. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A20: The effect of occupational following on job separations: robustness

	Same occ.	E2U job transition						Placebo test	
	First stage	OLS	IV					First stage	Reduced form
	(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)
Hiring rate	0.105***							0.006	-0.004
	(0.006)							(0.005)	(0.005)
Same occupation		0.004**	-0.096**	-0.098**	-0.088	-0.080			
		(0.002)	(0.047)	(0.047)	(0.058)	(0.061)			
Partial F-statistic	394		394	396	264	236		2	
First stage magnitude	0.13%								
Following rate	0.084	0.084	0.084	0.084	0.084	0.080		0.080	
Observations	362,878	362,878	362,878	361,183	360,165	362,826		353,591	353,591
Additional controls	None	None	None	Growth	None	Coarse		None	None
Time of hiring rate	Entry	Entry	Entry	Entry	Prior	Entry		Entry	Entry

Notes: All children included with respect to primary earning parent. Occupation taken at labor market entry. Two-digit occupation shown. Hiring rate shown in 10,000's. Demographic controls include age, sex, native born, language, and sex and income of primary earner. "Growth" time control specified as regional growth rate in hiring in occupation. "Prior" time of hiring rate refers to hiring rate in year prior to labor market entry. "Coarse" control includes hiring rate for one-digit coarser occupation in local labor market. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A21: Occupational and firm following, and firm AKM effects

	Log income	AKM effect
	(1)	(2)
Same occupation	0.020***	0.006***
	(0.001)	(0.001)
Same firm	0.015***	0.056***
	(0.002)	(0.002)
Observations	548,242	542,018
Occupational following rate	0.083	0.083
Firm following rate	0.045	0.045

Notes: OLS specification (Column (2) of Table 4) shown. Occupation and firm taken at labor market entry. Demographic controls include age, sex, native born, language, and sex and income of primary earner. AKM effect is firm AKM effect (Abowd et al. (1999)) in firm upon labor market entry. Standard errors clustered at primary earner level and shown in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A22: Summary of model parameters and identification

Model object	Parameter	Estimation and identification
Panel A. Education stage		
Skill transmission parameters	$(\mu_{kj}, \sigma^2_{kj})$	Pre-estimated for each skill k
General college cost κ^c	ϕ^c_0,ϕ^c_1,ϕ^c_{2k},ϕ^c_3	Estimated in model; identified by differences in college completion, conditional on individual-specific major values
Specific college major cost κ^m	ϕ^m_1	Estimated in model; identified by same major share
Panel B. Labor market stage		
Sampling distribution	$\Upsilon(\mathbf{y}\|m)$	Pre-estimated from the data (see Section H.1 for more details)
Job destruction rate	δ_0	Pre-estimated as sample mean E2U rate (see Section H.1 for more details)
Job offer arrival rates	λ_0, λ_1	Estimated in model; identified from U2E and E2E transition rates
Unemployment income	b	Estimated in model; identified from sets of acceptable jobs
Production function f	α_T,α^x_k,α^y_k,π_1	Estimated in model; identified from wages
Utility function c	π_2	Estimated in model; identified from occupational persistence rate
Skill accumulation function h	γ_k, g	Estimated in model; identified from correlations between skills and occupational skill requirements

Notes: Throughout the table, k refers to the different skill dimensions (q, v, s, m). Table refers to moments that are targeted to help identify each parameter.

Table A23: Targeted moments: data and model values

Moment	Data	Model
College share (parent college-educated)	0.401	0.381
College share (parent not college-educated)	0.612	0.585
College share (top half of x_q distribution)	0.607	0.630
College share (bottom half of x_q distribution)	0.272	0.277
College share (top half of x_v distribution)	0.607	0.624
College share (bottom half of x_v distribution)	0.280	0.284
College share (top half of x_s distribution)	0.514	0.523
College share (bottom half of x_s distribution)	0.380	0.384
College share (top half of x_m distribution)	0.411	0.417
College share (bottom half of x_m distribution)	0.481	0.490
Share with major that leads to parent occupation	0.099	0.098
U2E transition rate	0.236	0.247
E2E transition rate	0.172	0.131
Wage regression (x_{q0} coefficient)	0.036	0.023
Wage regression (x_{v0} coefficient)	-0.007	0.001
Wage regression (x_{s0} coefficient)	0.034	0.005
Wage regression (x_{m0} coefficient)	0.017	0.008
Wage regression (y_q coefficient)	0.565	0.383
Wage regression (y_v coefficient)	-0.334	-0.346
Wage regression (y_s coefficient)	0.028	0.018
Wage regression (y_m coefficient)	0.107	0.007
Wage regression (age coefficient)	0.038	0.043
Wage regression (tenure coefficient)	0.032	0.011
Wage regression (constant)	8.945	9.772
Mean wage at age 25	9.161	9.705
Mean wage at age 32	10.179	10.19
Mean wage at age 40	10.424	10.558
corr(x_{q0},y_q) at age 32	0.46	0.418
corr(x_{v0},y_v) at age 32	0.476	0.558
corr(x_{s0},y_s) at age 32	0.239	0.211
corr(x_{m0},y_m) at age 32	0.207	0.06
corr(x_{q0},y_q) at age 40	0.462	0.512
corr(x_{v0},y_v) at age 40	0.478	0.626
corr(x_{s0},y_s) at age 40	0.256	0.277
corr(x_{m0},y_m) at age 40	0.22	0.122
Fraction in same occupation	0.111	0.114
Same occupation wage premiuim	0.023	0.023

Notes: Table shows data and model values of all targeted moments.

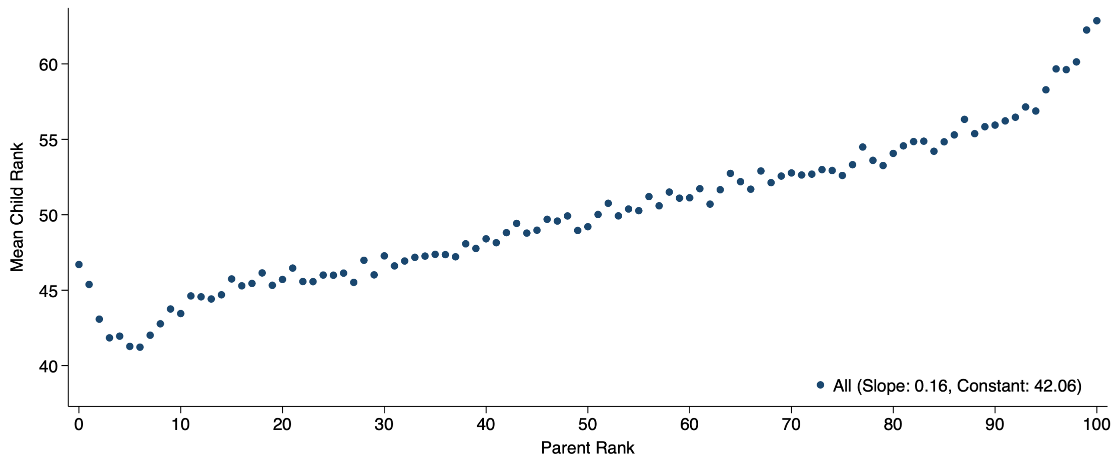
Table A24: Model-based decomposition of occupational following: mechanism interactions

	Rate of following			Total output
	Class	1-Digit	2-Digit	
	(1)	(2)	(3)	(4)
Baseline	0.537	0.246	0.138	
Equalize skill transmission	-99.2%	-16.5%	-9.0%	0.1%
	[0.5]	[0.224]	[0.128]	
Equalize college cost	-99.4%	-16.5%	-9.1%	0.1%
	[0.5]	[0.224]	[0.128]	
No major cost	-101.1%	-17.6%	-12.6%	0.12%
	[0.5]	[0.222]	[0.124]	
No idiosyncratic preferences	-113.0%	-23.6%	-20.8%	0.11%
	[0.495]	[0.214]	[0.115]	
No differential search frictions	-113.8%	-24.2%	-21.7%	0.11%
	[0.495]	[0.213]	[0.114]	
No occupation-specific skill	-131.1%	-51.2%	-64.1%	-0.07%
	[0.489]	[0.177]	[0.066]	
Equalize education stage	-100.9%	-17.6%	-12.6%	0.12%
	[0.5]	[0.222]	[0.124]	
Equalize labor stage	-31.4%	-37.5%	-57.7%	-0.23%
	[0.525]	[0.195]	[0.073]	
Equalize all	-130.8%	-51.2%	-64.3%	-0.07%
	[0.489]	[0.177]	[0.066]	

Notes: Percentage change in rate of occupational following from model-implied baseline case shown, with overall level shown in brackets. Percentage change is relative to random sorting benchmark, as defined in Appendix Section H.4. I compute the percentage change as $1 - \frac{\text{simulated rate}-\text{random benchmark}}{\text{baseline rate}-\text{random benchmark}}$. Shown for all children. Shown at class level, 1-digit level, and 2-digit level. Percentage change in total output from model-implied baseline case shown. All mechanisms considered jointly, in chronological order (i.e., mechanisms considered along with interactions).

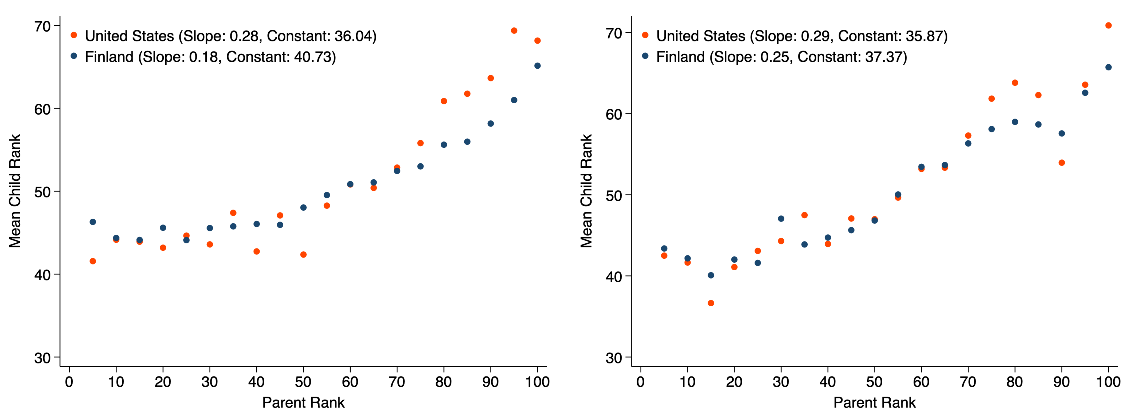
B Appendix Figures

Figure A1: Intergenerational income mobility across the population



Notes: All children included, with respect to total parental income. Using mean child income from ages 25-35, and then ranked within birth cohort. Only children for which I observe until age 35 included, and so therefore limited to cohorts born between 1975 and 1983. Total parental income taken as average income over ages 25-55, and then ranked within child birth cohort. Constant defined at parental income rank 0.

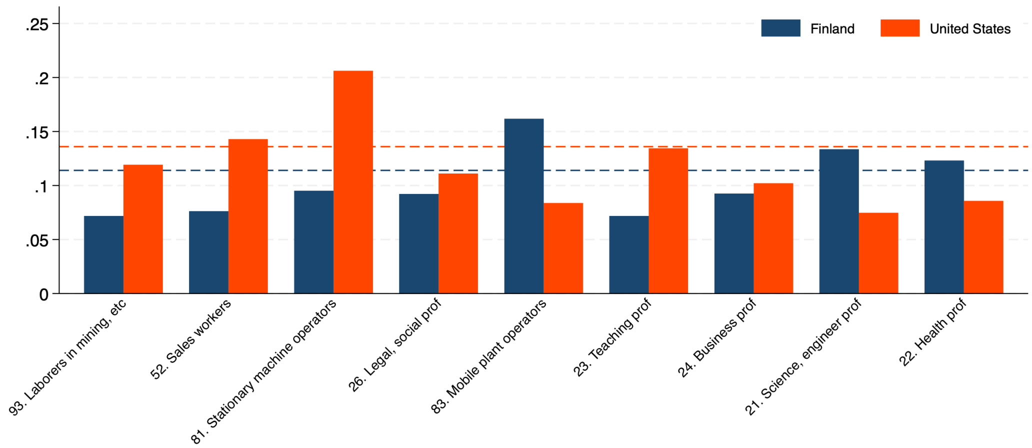
Figure A2: Income and occupational mobility: Finland vs. United States



(a) Intergenerational income mobility (b) Intergenerational occupational mobility

Notes: Data for Finland taken from my sample. Data for United States comes from the Panel Study of Income Dynamics (PSID). Shown at age 35 for males with respect to their fathers, due to PSID data limitations. Child ranked within cohort. Parent ranked within child cohort. Occupational rank taken as mean income in two-digit occupation.

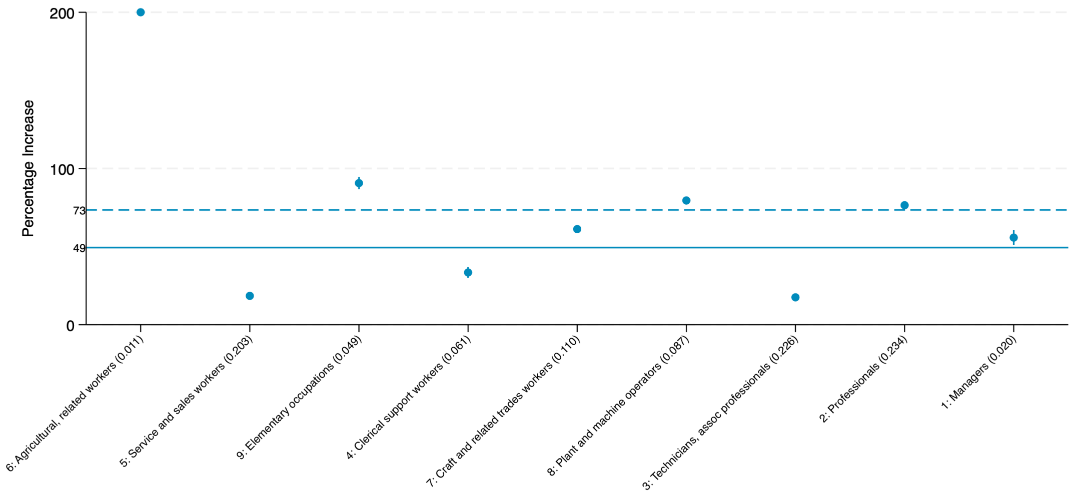
Figure A3: Occupational following: Finland vs. United States



Notes: Data from Finland taken from my sample. Data for United States taken from the Panel Study of Income Dynamics (PSID). Shown at two-digit level, for occupations for which there is an unambiguous mapping from U.S. Census classification to ISCO classification. Shown at age 30 for males with respect to their fathers, due to PSID data limitations. Occupations are ordered by mean income in occupation.

Figure A4: Occupational persistence across the distribution

(a) One-digit level



(b) Two-digit level

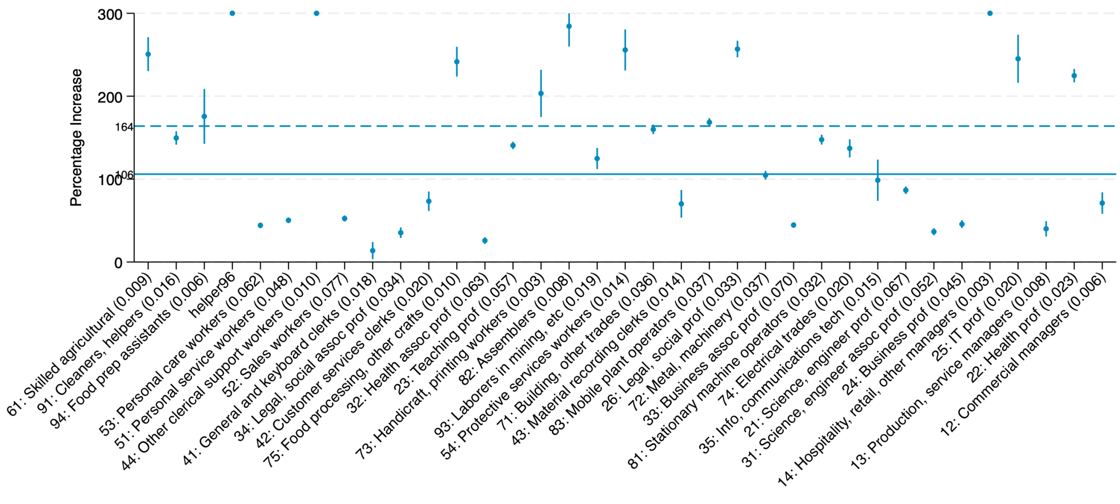
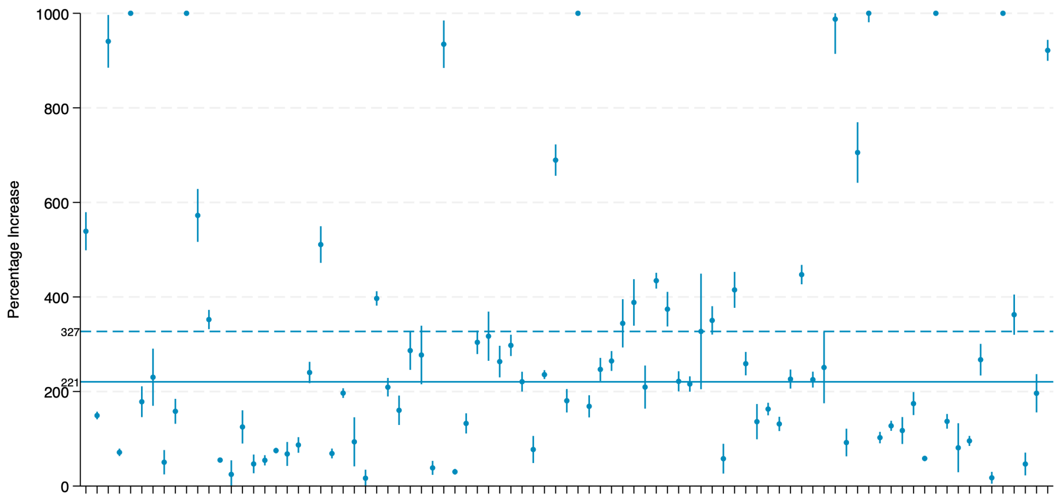
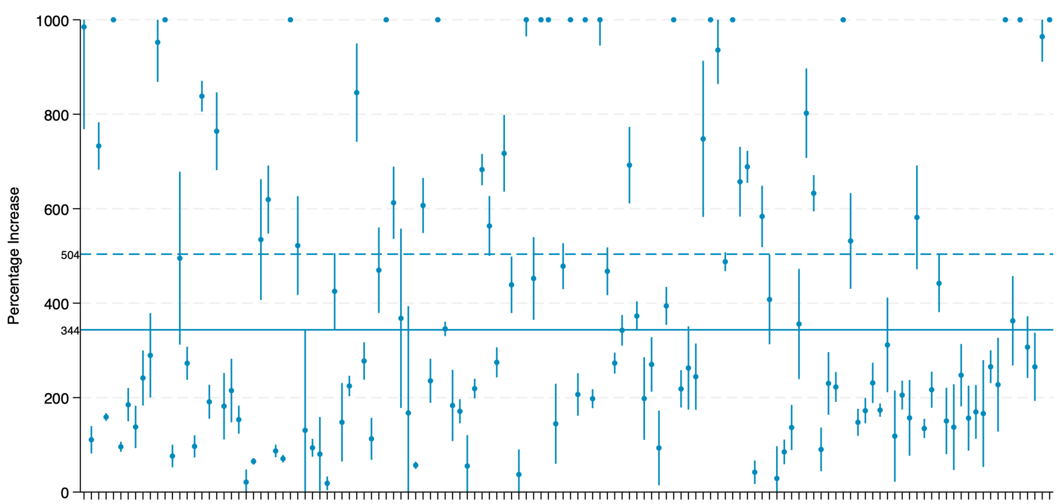


Figure A4: Occupational persistence across the distribution (cont.)

(c) Three-digit level

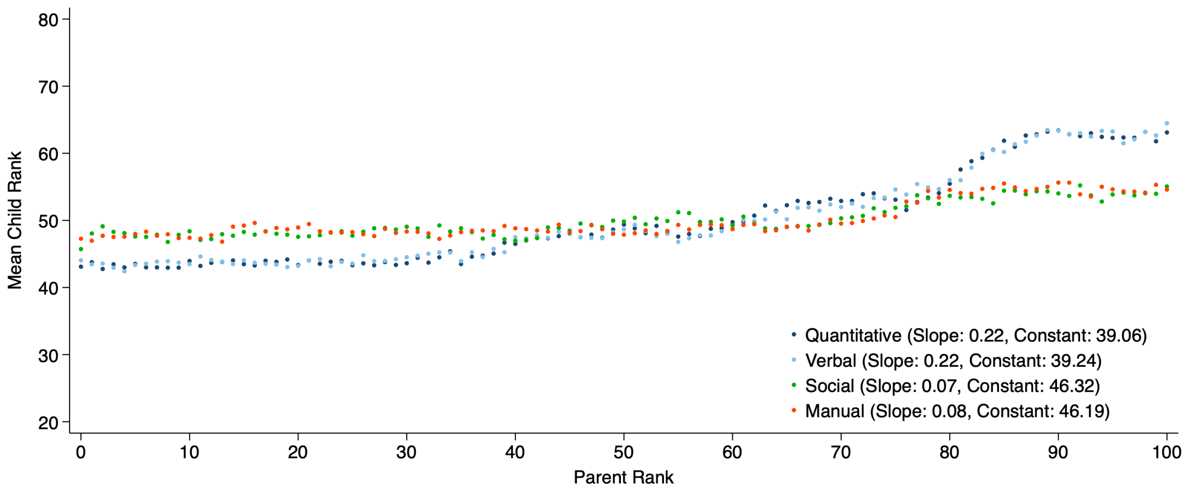


(d) Four-digit level



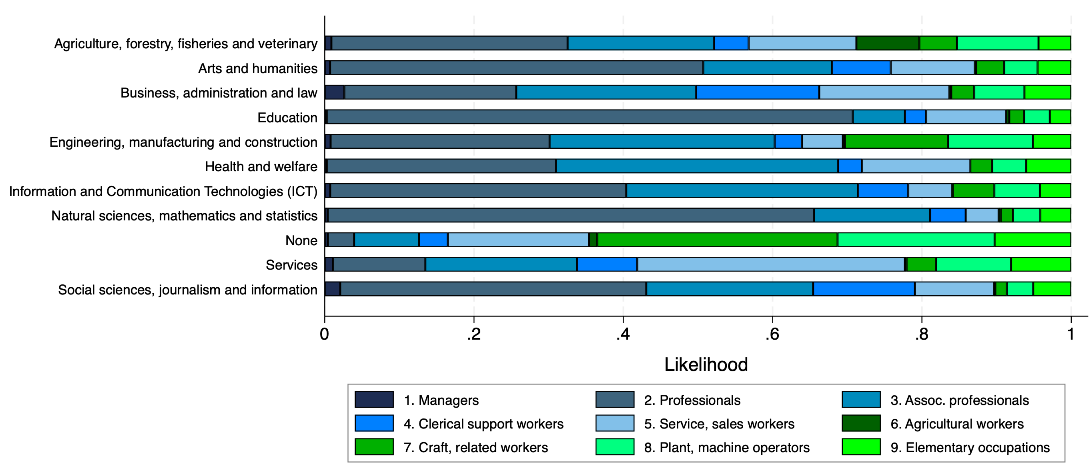
Notes: All children included, with respect to union of both parents. Occupations ordered in increasing order of occupation-specific mean income. Solid line shows weighted average while dashed line shows unweighted average across all occupations, and corresponding values labeled on vertical axis. Only the occupations that are more than 0.2% of the population are shown and figures truncated at top, each for clarity of illustration.

Figure A5: Intergenerational persistence in skills: all implied skills



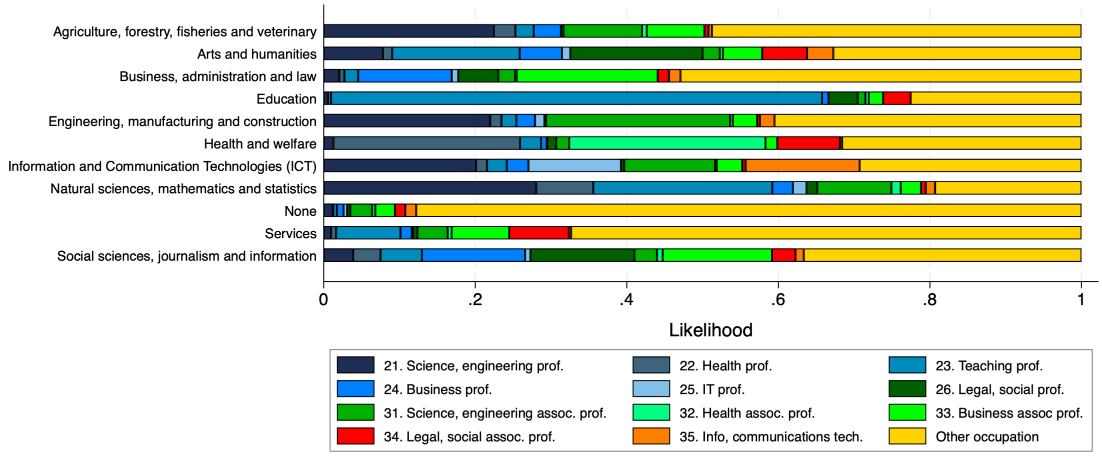
Notes: Data only available for males. Father cognitive skills and personality traits implied by their most-experienced 2-digit occupation. Skills are standardized individually within cohort. Child rank defined within cohort. Parent rank defined within child cohort.

Figure A6: Occupational sampling distribution: one-digit level (aggregated occupations)



Notes: Figure shows stacked likelihood of receiving an offer from each occupation, conditional on college major. The width of each bar represents the probability of receiving an offer from that occupation, conditional on one's college major. Aggregated from two-digit occupational level to one-digit occupational level.

Figure A7: Occupational-sampling distribution: two-digit level (select occupations)



Notes: Figure shows stacked likelihood of receiving an offer from each occupation, conditional on college major. The width of each bar represents the probability of receiving an offer from that occupation, conditional on one's college major. Shown at two-digit occupational level, only for one-digit occupations 2 (professionals) and 3 (associate professionals), with "Other" category including all other two-digit occupations.

C Data description and variable definitions

In this appendix section, I describe the data and key variable definitions. The demographic data and some education data comes from the FOLK Basic data module provided by Statistics Finland. The college major comes from the FOLK Degree module. The labor market data primarily comes from various FOLK modules, including FOLK Income, FOLK Employment, and FOLK Unemployment. The occupation variable comes primarily from the `ammattikoodi` variable in FOLK Employment (I discuss it more in Section C.2).

- Key demographic variables include the following:
  - Age
  - Sex
  - Country of birth: indicator if born in Finland
  - Language: Finnish, Swedish, or other
  - Municipality of residence: one of 472 different municipalities in Finland.
- I discuss the skill data in Appendix C.1.
- Key education variables include the following:
  - Educational level: one of 7 different categories (upper secondary level, post-secondary non-tertiary education, short-cycle tertiary education, bachelor's or equivalent level, master's or equivalent level, doctoral or equivalent level, other)
  - Field of study: as defined by the four-digit code from the International Standard Classification of Education (2013), which consists of 101 fields of study, such as economics, fine arts, and medicine
  - Field of study (broad): I also use the broad classification from the International Standard Classification of Education (2013), which consists of the following 11 broad field of study categories: generic; education; arts and humanities; social sciences, journalism and information; business, administration and law; natural sciences, mathematics and statistics; information and communications technologies; engineering, manufacturing and construction; agriculture, forestry, fisheries and veterinary; health and welfare; services; unknown
- Key labor market variables include the following:
  - Income: total earned income

– Firm: the establishment for the last employment relationship in the year (`sykstun` from the FOLK TKT data set)

- Finally, other remaining key variables include the following:

  – The FOLK basic data includes the main activity for each individual each year, which can be employment, unemployment, child, student, pensioner, conscript, on unemployment pension, or outside the labor force.

C.1 Construction of skill measures

In this section, I discuss how I construct worker skill measures. As discussed in Section 2.2, I have from the Finnish Defence Forces measures of cognitive skills for individuals born between 1964 and 1997 and personality measures for individuals born between 1964 and 1982, all for most males. I construct the multidimensional skills of workers at age 18, after they typically complete all secondary education and when they decide whether or not to attend tertiary education (and if so, what field of study to pursue). I use the following measures to construct the multidimensional skill bundle for workers at age 18. I use the following 11 measures from the Defence Forces, which include three measures related to cognitive skills and eight measures related to personality traits. The following are short descriptions of each measure, taken from Jokela et al. (2017):

- Cognitive ability:

  – Arithmetic: "completing a series of numbers that follow a certain pattern, solving short verbal problems, computing simple arithmetic operations, and choosing similar relationships between two pairs of numbers."

  – Verbal: "choosing synonyms or antonyms of a given word, selecting a word that belongs to the same category as a given word pair, choosing which word on a list does not belong in the group, and choosing similar relationships between two word pairs."

  – Visuospatial: "set of matrices containing a pattern problem with one removed part, and the participant needs to decide which of the given alternative figures completes the matrix; it is similar to Raven's Progressive Matrices."

- Personality traits:

- Achievement striving: "how strongly the person wants to perform well and achieve important life goals (24 items; e.g., whether the person is prepared to make personal sacrifices to achieve success)."
- Activity-energy: "how much the person exerts physical effort in everyday activities and how quickly the person prefers to execute activities (28 items; e.g., whether the person tends to work fast
- Deliberation: "how much the person prefers to think ahead and plan things before acting (26 items; e.g., whether the person prefers to spend money carefully)"
- Dutifulness: "how closely the person follows social norms and considers them to be important (18 items; e.g., whether the person would return money if given back too much change at a store)."
- Leadership motivation: "how much the person prefers to take charge in groups and influence other people; it includes 30 items."
- Masculinity: "the person's occupational and recreational interests that are traditionally considered as masculine (27 items; e.g., whether the person would like to work as a construction manager)."
- Self-confidence: "the person's self-esteem and beliefs about his abilities (32 items; e.g., whether the person feels to be as good and able as others and can meet other people's expectations)."
- Sociability: "the person's level of gregariousness and preference for socializing with others (33 items; e.g., whether the person likes to host parties and not withdraw from social events)." and vigorously and prefers fast-paced work)."

I condense the measures into four main skills with the following exclusion restrictions:

1. Quantitative: arithmetic measure
2. Verbal: verbal measure
3. Social: sociability measure
4. Manual: masculinity measure

I condense the measures following Lise and Postel-Vinay (2020) by conducting these four steps:

1. Standardize measures and data structure: I standardize measures to have mean zero and standard deviation one within each cohort. The data includes P different measures for N individuals, which I denote by the $N \times P$ matrix $\mathbf{M}$.

2. Principal component analysis: I use Principal Component Analysis (PCA) to decompose the matrix $\mathbf{M} = \mathbf{FL}$, where $\mathbf{F}$ is an orthonormal $N \times P$ matrix of principal eigenvectors of $\mathbf{M}^\top\mathbf{M}$ and $\mathbf{L}$ is a $P \times P$ matrix of factor loadings. I take the first four principal components and consider the decomposition $\mathbf{M} = \mathbf{F}_4\mathbf{L}_4 + \mathbf{U}$, where $\mathbf{F}_4$ is the $N \times 4$ matrix formed by taking the first four columns of $\mathbf{F}$ and $\mathbf{L}_4$ is the $4 \times P$ matrix formed by taking the first four rows of $\mathbf{L}$.

3. Interpretability and exclusion restrictions: I then use exclusion restrictions to transform the decomposition into interpretable skills. For any 4×4 invertible matrix $\mathbf{T}$, the above decomposition of $\mathbf{M}$ can be written as $\mathbf{M} = (\mathbf{F}_4\mathbf{T})(\mathbf{T}^{-1}\mathbf{L}_4) + \mathbf{U}$, which is an alternative decomposition of the measures $\mathbf{M}$ into new factors $\mathbf{F}_4\mathbf{T}$ with loadings $\mathbf{T}^{-1}\mathbf{L}_4$. I choose $\mathbf{T}$ such that the final decomposition of $\mathbf{M}$ satisfies the chosen exclusion restrictions written above.

4. Rescale: Lastly, I rescale the four measures so that they lie in the interval $[0, 1]$.

C.2 Occupation data

I compute two occupation variables for each individual: their current occupation and their modal (i.e., most-experienced) occupation at that point. I use the following data to construct this:

1. I use FOLK administrative data that covers 100% of the population for 1990, 1993, 1995, 2000, and 2004-present.

2. I then fill in the gaps in time with the occupation on either side if:

   (a) The individual has the same occupation before and after.

   (b) The individual is in the same firm before and after.

   (c) Their income changes by less than 20% over the time period.

3. I lastly fill in remaining gaps with the occupation variable from the Structure of Earnings Statistics (SES) data set, which is an annual survey starting in 1995 that covers a large sample of the population.[48]

[48] The SES includes all public sector workers and a sample of about 55-75% of private sector workers.

I use the occupational classification from the International Classification of Occupations (ISCO; International Labour Organization (2023)). There are 9 one-digit codes, 43 two-digit codes, 130 three-digit codes, and 436 four-digit codes. Table A25 illustrates the one-digit codes and their corresponding skill levels according to the ISCO. I also show which occupations I classify as white collar occupations as opposed to blue collar occupations, directly based on differences in their skill levels. Also note that, as discussed in Section 2.3, I drop all armed forces occupations from the sample (one-digit code 0).

Table A25: One-digit occupational codes and white collar definition

Occupation	ISCO skill level	WC?
1. Managers	3,4	✓
2. Professionals	4	✓
3. Technicians and assoc. professionals	3	✓
4. Clerical support workers	2	
5. Service and sales workers	2	
6. Agricultural, forestry, fishery workers	2	
7. Craft and related trades workers	2	
8. Plant and machine operators	2	
9. Elementary occupations	2	
0. Armed forces	1,2,4	

Notes: The table illustrates all one-digit codes, their corresponding skill level as given by the ISCO, and whether or not I classify them as a white collar occupation.

Table A26 provides two examples of the occupational classification. Example 1 starts with the one-digit code 2, which are professionals and which includes six two-digit codes. Two-digit code 22, health professionals, then includes six three-digit codes. Finally, three-digit code 221, medical doctors, includes two four-digit codes: generalist medical practitioners (2221) and specialist medical practitioners (2222). This example illustrates how each category becomes progressively more specific. Example 2 illustrates this for a set of blue collar occupations.

Table A26: Examples of occupational classification

Example 1:	Example 2:
2. Professionals include...	7. Craft and related trades workers include...
21. Science, engineering prof.	71. Building, other trades
22. Health prof.	72. Metal, machinery
23. Teaching prof.	73. Handicraft, printing workers
24. Business prof.	74. Electrical trades
25. IT prof.	75. Food processing, other crafts
26. Legal, social prof.	
22. Health professionals include...	71. Building, other trades include...
221. Medical doctors	711. Building frame workers
222. Nursing prof.	712. Building finishers
223. Traditional medical prof.	713. Painters, cleaners, etc.
224. Paramedical practitioners	
225. Veterinarians	
226. Other health prof.	
221. Medical doctors include...	711. Building frame workers include...
2221. Generalist medical practitioners	7111. House builders
2222. Specialist medical practitioners	7112. Bricklayers and related workers
	7113. Stonemasons, stone cutters, etc.
	7114. Concrete placers, finishers, etc.
	7115. Carpenters and joiners
	7119. Other

Notes: The table provides two examples of the occupational classification. It starts with a one-digit occupation, and then progressively includes all the finer occupations that are subsumed by it up until the four-digit level.

C.3 Key variable definitions

I define occupational followers and firm followers as:

- Occupational follower: whether the child's modal or current occupation is the same as the parent's modal occupation.
- Firm follower: whether the child's firm is the same as the parent's firm.

I also compute two measures of skill mismatch, following Fredriksson et al. (2018); Guvenen et al. (2020); Lise and Postel-Vinay (2020). First, I consider horizontal mismatch, which captures the degree to which the specific skills of a worker align with the specific skill requirements of their occupation. In particular, I define this as:

$$\text{H. Mismatch}_{ij} = \sum_{k=q,v,s,m} |x_{ik} - y_{jk}|/4, \tag{10}$$

where x_{ik} is the individual i's amount of skill in dimension k and y_{jk} is the skill requirement of occupation j along the same skill dimension k, as defined in Section 2.4. I then standardize this measure in the sample to have mean zero and standard deviation one.

Second, I consider a measure of vertical mismatch, which captures the degree to which the average overall skill level of a worker aligns with the average overall skill requirement of their occupation. This is defined as:

$$\text{V. Mismatch}_{ij} = |\sum_{k=q,v,s,m} x_{ik}/4 - \sum_{k=q,v,s,m} y_{jk}/4|, \tag{11}$$

where $\sum_{k=q,v,s,m} x_{ik}/4$ is the average skill level of individual i and $\sum_{k=q,v,s,m} y_{jk}/4$ is the average skill requirement of occupation j. Again, I then standardize this measure in the sample to have mean zero and standard deviation one.

D The relationship between occupational persistence and income persistence

In this section, I show that occupational persistence is (a) associated with less upward income mobility and (b) can explain a substantial share of the population-wide level of income persistence.

D.1 From the individual's perspective

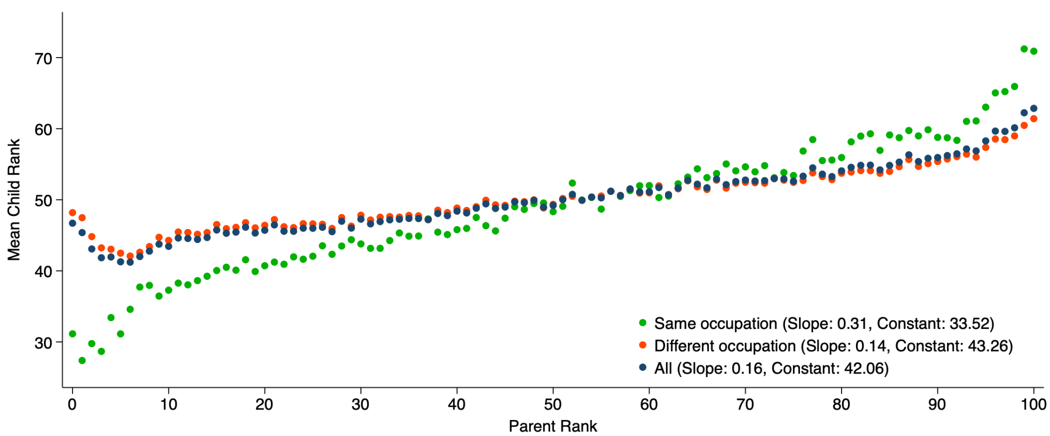
I begin by estimating to what degree occupational following—from the individual's perspective—is associated with income persistence. I first estimate the following regression equation, which is the standard rank-rank regression in income:

$$R_c = \beta_0 + \beta_1 \cdot R_p + \varepsilon, \tag{12}$$

where R_c is the income rank of the child and R_p is the income rank of the parent. The main coefficient of interest is β_1, which captures the degree to which income persists across generations. To then consider the degree to which occupational following is associated with income persistence, I estimate this equation separately for children that enter their parent's occupation and for children that do not.

I illustrate the results in Figure A8. The overall rank-rank slope between parent income rank and child income rank in the sample is 0.16, which is similar to other estimates in Finland (e.g., Kaila et al. (2024)). I then show that this relationship varies greatly by whether or not the child is in the same occupation as their parent. First, children that enter their parent' occupation experience less relative mobility than those that do not: the coefficient on parent income rank more than doubles from 0.14 to 0.31. Second, children that enter their parent's occupation experience less absolute mobility: for children born to parents at the 25th percentile of the income distribution, the average income rank for children that enter their parent's occupation is 41.3, while it is 46.8 for those that do not, or 5.5 ranks larger. Taken together, this means that occupational following is associated with significantly less upward income mobility.

Figure A8: Intergenerational income mobility and occupational following



Notes: All children included, with respect to total parental income. Using mean child income from ages 25-35, and then ranked within birth cohort. Only children for which I observe until age 35 included, and so therefore limited to cohorts born between 1975 and 1983. Current two-digit occupation at age 35 used. Total parental income taken as average income over ages 25-55, and then ranked within child birth cohort. Constant defined at parental income rank 0. Same occupation defined as if in same occupation as either parent.

D.2 From the population's perspective

I now estimate to what degree occupational choice and occupational persistence can explain the overall population-wide level of income persistence. Given the fact that I largely follow Haeck and Laliberté (2024) and the computational complexity in conducting this exercise, I leave the estimation details to Appendix Section E. First, as shown in Table A27, I find that the sorting of children into occupations on the basis of parental income can explain 37-60%

(depending on the occupational classification) of the rank-rank relationship in income. This is a large amount—as a point of comparison, I show in Appendix Table A28 the degree to which other observable characteristics (such as skills and education) can explain this relationship, and find that none can explain as much of income persistence as occupational choice can. I then estimate how much income persistence can be explained by the complete intergenerational occupational transition matrix—which characterizes the entire distribution of occupations that children pursue given their parental occupation. I find that this can explain 26-46%. Finally, I go one step further than Haeck and Laliberté (2024) and estimate how much income persistence can be explained by occupational following—when children enter into the *exact same* occupation as their parents. I find that occupational following can explain 5-13%.

Table A27: Occupational choice and population-wide income persistence

	Digit level			
	1	2	3	4
Occupational choice explains:	36.7	49.6	54.6	60.1
Occupational transmission explains:	26.3	34.9	40.1	46.0
Occupational following explains:	12.5	5.3	7.4	6.5

Notes: Table shows the share of the population-wide level of income persistence that can be explained by each factor, by occupational clasification (digit level). Occupational choice refers to differential sorting of children into occupation by parental income. Occupational transmission refers to the entire intergenerational occupational transition matrix. Occupational following refers to children choosing the exact same occupation as their parents. All children included, with respect to union of parents. See Appendix Section E for more details.

Taken together, this is evidence that the way children choose occupations, and how they are influenced by their parents, can explain a quantitatively large share of the overall level of income persistence in the population.[49] I view both the role of occupational following in particular and occupational choice in general to be relevant for this paper. While the first half of this paper pays special attention to occupational followers, the second half embeds occupational following into a more general occupational choice model. Thus, the degree to which parents affect the occupational choices of their children, above and beyond just

[49] These shares are comparable to, and sometimes larger than, the relative importance of other key economic factors in shaping intergenerational mobility, such as neighborhoods (Chetty and Hendren (2018a,b)), migration (Anstreicher (2024)), labor market shocks (Kaila et al. (2024)), and firms (Dobbin and Zohar (2024); Engzell and Wilmers (2024); Staiger (2023)). For example, Staiger (2023) finds that if children were not allowed to enter into their parent's firm, the relationship between initial earnings and parent earnings would be 7% lower, while Kaila et al. (2024) finds that the differential impacts of job loss can explain 3.7% of the population-wide level of income persistence.

Table A28: Intergenerational income mobility and its mediators

	Child income rank					
	(1)	(2)	(3)	(4)	(5)	(6)
Parent income rank	0.191***	0.151***	0.094***	0.151***	0.118***	0.078***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Same occupation		-18.75***				
		(0.597)				
Parent income rank · same occupation		0.333***				
		(0.010)				
Constant	40.44***	42.71***	45.31***	41.28***	44.10***	45.50***
	(0.196)	(0.207)	(0.178)	(0.192)	(0.189)	(0.175)
Observations	83,960	83,960	83,960	83,960	83,960	83,960
Controls	None	None	Occupation	Skills	Education	All

Notes: Sample limited to Defence Forces sample, which are only males, as this exercise requires skill data. Taking with respect to fathers. Using mean child income from ages 25-35, and then ranked within birth cohort. Current two-digit occupation at age 35 used. Skills control includes linear control for each skill individually. Education control includes fixed effects for amount and field of education (using specific field of study, which has 101 categories; see Appendix Section C for more details). Total father income taken as average income over ages 25-55, and then ranked within child birth cohort. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

focusing on those that enter into their parent's exact occupation, is relevant.

E Intergenerational income persistence decomposition

I start by following Haeck and Laliberté (2024) and estimating a version of the rank-rank specification from Equation 12 with child occupation fixed effects. As shown in Column (2) of Table A29, the rank-rank slope, conditional on child occupational choice, is 0.073. This implies that nearly half, or 46% ($= 1 - \frac{0.073}{0.136}$), of the unconditional rank-rank relationship in income is due to the differential occupational choices of children by parental income rank. This is a large amount—as a point of comparison, I show in Appendix Table A28 the degree to which other observable characteristics (such as skills and education) can explain this relationship, and find that none can explain as much of income persistence as occupational choice can. Moreover, this is in the range of what Haeck and Laliberté (2024) find: they use Canadian data and find that occupational choice explains 20-50% of intergenerational income persistence.

Table A29: Intergenerational income mobility and occupational choice

	Child income rank	
	(1)	(2)
Parent income rank	0.136***	0.073***
	(0.001)	(0.001)
Constant	43.21***	46.35***
	(0.070)	(0.063)
Observations	675,264	675,264
Occupation FE		Y

Notes: All children included, with respect to total parental income. Using mean child income from ages 25-35, and then ranked within birth cohort. Only children for which I observe until age 35 included, and so therefore limited to cohorts born between 1975 and 1983. Current two-digit occupation at age 35 used. Total parental income taken as average income over ages 25-55, and then ranked within child birth cohort. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

I then further follow Haeck and Laliberté (2024) to quantify the degree to which the intergenerational transmission of occupation—how parental occupation affects the child's occupation—can explain this persistence, rather than how parental income more broadly affects child occupational choice and child income. In essence, I compute counterfactual occupational choices where I assume that the intergenerational transmission of occupation is constant across all parental income groups, and then estimate the occupational returns by parental income group. Finally, I go one step further than Haeck and Laliberté (2024) and estimate to what degree occupational following—when children enter into the *exact same* occupation as their parents—can explain the overall level of intergenerational income persistence. I similarly do this by computing counterfactual occupational choices where I assume that occupational following, by occupation, is constant across all parental income groups.

I plot the results in Appendix Figure A9. At the two-digit level, general intergenerational occupational transmission can explain 69.8% of the overall sorting and thus 32.4% of the overall level of income persistence.[50] Specific occupational following can explain 12.7% of the overall sorting and thus 5.9% of the overall level of income persistence.

E.1 Role of occupational transmission in general

In this section, I discuss how I quantify the degree to which occupational choice can explain the intergenerational rank-rank relationship in income. I follow Haeck and Laliberté (2024), with a few modifications, and estimate the following unconditional rank-rank relationship:

$$R_c = \sum_{P=1}^{100} b_i \mathbb{1}\{R_p = i\} + \epsilon, \tag{13}$$

where R_c is the child's income rank and R_p is their parent's income rank.

I next follow Haeck and Laliberté (2024) and estimate a conditional rank-rank relationship in income, where I control for child occupation fixed effects, by estimating the following equation:

$$R_c = \sum_{P=1}^{100} \beta_P \mathbb{1}\{R_p = i\} + \sum_{j \in \mathcal{J}} \delta_{j(c)} + \varepsilon, \tag{14}$$

where R_c is again the child's income rank, R_p is again their parent's income rank, and $\delta_{j(c)}$ is an occupation fixed effect for $j \in \mathcal{J}$ for the occupation of the child $j(c)$.

Table A29 illustrates the unconditional and conditional estimation results—assuming a linear slope—in Columns (1) and (2), respectively. Note that the slope of the unconditional relationship is 0.136, while the slope of the conditional relationship is 0.073. This implies that controlling for child occupation explains 46% ($=1 - \frac{0.073}{0.136}$) of the overall rank-rank

[50] This estimate is similar to, although slightly larger than, the estimate from Haeck and Laliberté (2024), whose estimate is around 20% when occupation is measured at the same time as my estimate.

relationship. To be precise, this means that the way in which children differentially choose occupation given their parental income explains 46% of the overall rank-rank relationship. In order to estimate how the transmission of parental occupation to child occupation affects this relationship, one must do something more involved, as I do next.

Following Haeck and Laliberté (2024), the average income rank for a child with parental income rank P can be written as follows:

$$\bar{y}_P = E[R_c|R_p = P] = \sum_{j\in\mathcal{J}} \bar{R}_{j|P} s_{j|P} = \underbrace{\sum_{j\in\mathcal{J}} (\bar{y}_{j|P} - \hat{\delta}_j) s_{j|P}}_{\hat{\beta}_P} + \underbrace{\sum_{j\in\mathcal{J}} \hat{\delta}_j s_{j|P}}_{\Delta_P}, \tag{15}$$

where $\bar{y}_{j|P}$ is the average income rank of children in occupation j with parental income rank P, $s_{j|P}$ is the share of children with parental income rank P that choose occupation j, and $\hat{\delta}_j$ are the estimated returns to occupation j from Equation 14 above. Here, Δ_p is the amount of income persistence that can be explained by the sorting of children into occupations differentially by parental income rank.

Next, I estimate the degree to which occupational transmission can explain overall income persistence, again following Haeck and Laliberté (2024). Let $\nu_{q|P}$ be the share of parents with income rank P that are in occupation q. Also, let $s_{j|P,q}$ be the share of children with parents with income rank P and from occupation q that choose occupation j. Note that the only difference from $s_{j|P}$ is that I also condition on the parent's occupation q. These two quantities are related via $s_{j|P} = \sum_{q\in\mathcal{J}} s_{j|P,q}\nu_{q|P}$.

I then can write the term Δ_P as

$$\Delta_P = \sum_{j\in\mathcal{J}} \hat{\delta}_j s_{j|P} = \sum_{j\in\mathcal{J}} \hat{\delta}_j (\sum_{q\in\mathcal{J}} s_{j|P,q}\nu_{q|P}) = \sum_{j\in\mathcal{J}} \hat{\delta}_j \sum_{q\in\mathcal{J}} (s_{j|P,q} - s_{j|q})\nu_{q|p} + \underbrace{\sum_{j\in\mathcal{J}} \hat{\delta}_j \sum_{q\in\mathcal{J}} s_{j|q}\nu_{q|P}}_{\text{Differences in parental occ.}}, \tag{16}$$

where now $s_{j|q}$ is the share of children with parental occupation q across all income ranks that choose occupation j. I estimate the second term, which represents the role of intergenerational occupational transmission by restricting the occupational transmission to be the same across all parental income groups.

E.2 Role of occupational following in particular

I now finally consider the role of occupational following—when children enter into the exact same occupation as their parent—in particular. To do so, I consider the following regression-based approach from Haeck and Laliberté (2024), which is equivalent to their re-weighting approach, as just described.

Consider the following regression for all occupations $j \in \mathcal{J}$:

$$\mathbb{1}\{j(i) = j\} = \mu_q^j + \rho_{qj}^o, \tag{17}$$

where $\mathbb{1}\{j(i) = j\}$ is an indicator for child i being in occupation j and μ_q^j are parental occupation fixed effects.

The fitted values for a given parental income rank P is equal to the following:

$$\tilde{s}_{j|P} = E[\mathbb{1}\{j(i) = j\}] = E[\mu_q^j|p] + [\rho_{qj}^o|p], \tag{18}$$

where $\tilde{s}_{j|P}$ is the estimated share of children from parental income rank P in occupation j. I then use this estimated share $\tilde{s}_{j|P}$ to compute the estimated parent rank-specific total occupational returns $\tilde{\Delta}_P = \sum_{j\in\mathcal{J}} \hat{\delta}_j \tilde{s}_{j|P}$. The comparison of $\tilde{\Delta}_P$ to Δ_P is then the contribution of occupational transmission to the role of sorting on occupation by parental income rank.

To then estimate the contribution of occupational following, I conduct a similar exercise, but limit the child's occupation to only be a function of whether their parent is in that occupation. In particular, I estimate the following regression equation for all occupations $j \in \mathcal{J}$:

$$\mathbb{1}\{j(i) = j\} = \xi \cdot \mathbb{1}\{q(i) = q\} + \varrho_{qj}^o, \tag{19}$$

where now $\mathbb{1}\{q(i) = q\}$ is an indicator if the occupation of the parent of child i is $q(i)$.

I then compute the following fitted values for a given parental income rank p:

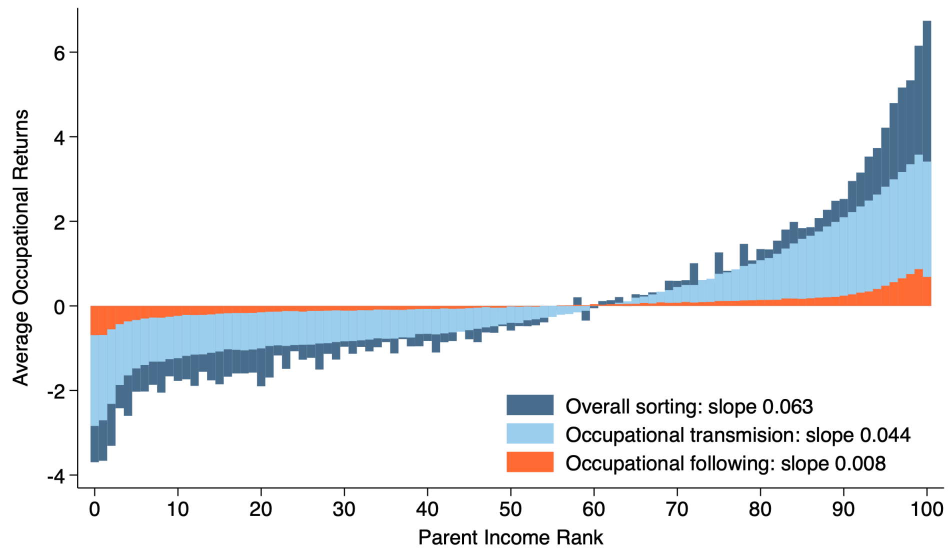
$$\tilde{\tilde{s}}_{j|P} = E[\mathbb{1}\{j(i) = j\}] = E[\xi \cdot \mathbb{1}\{q(i) = q\}|P] + [\varrho_{qj}^o|P]. \tag{20}$$

As before, I then use this estimated share to compute the estimated parent rank-specific total occupational returns $\tilde{\tilde{\Delta}}_P = \sum_{j\in\mathcal{J}} \hat{\delta}_j \tilde{\tilde{s}}_{j|P}$. The comparison of $\tilde{\tilde{\Delta}}_P$ to Δ_P is now the contribution of occupational following in particular to the role of sorting on occupation by parental income.

E.3 Decomposition results

Figure A9 illustrates the estimated Δ_p, $\tilde{\Delta}_p$, and $\tilde{\tilde{\Delta}}_p$. The slope of the overall term—which represents general sorting into occupation based on parental income—is 0.063. The slope of the term based on differences in overall occupational transmission is 0.044, which is 69.8% of the overall sorting term and thus 32.4% of the overall rank-rank slope. Finally, the slope of the term based on differences in occupational following is 0.008, which is 12.7% of the overall sorting term and thus 5.9% of the overall rank-rank slope.

Figure A9: Occupational following and overall income persistence



Notes: Figure shows average occupational returns as a function of parental income rank overall, assuming constant occupational transmission across all parent ranks, and assuming constant occupational following across all parent ranks. Shown at two-digit occupational level. Shown for all children, with respect to union of parents.

F Model structure details

The total private value of a match between a worker and a job is given by $P(\mathbf{x}, \mathbf{y})$. The value of a current job for a worker is given by $W(\mathbf{x}, \mathbf{y})$, while the value of unemployment for a worker is given by $U(\mathbf{x})$. It must be such that $U(\mathbf{x}) \leq W(\mathbf{x}, \mathbf{y}) \leq P(\mathbf{x}, \mathbf{y})$: the worker would quit their current job if the first inequality were not true, while the firm would terminate the worker if the second inequality were not true. I assume free entry of jobs, so that the value of a job vacancy is zero, thus giving the total surplus of a match as $P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})$. The worker's share of surplus σ is endogenous and is given by $\sigma = \frac{W(\mathbf{x},\mathbf{y}) - U(\mathbf{x})}{P(\mathbf{x},\mathbf{y}) - U(\mathbf{x})}$ while the firm's share of surplus is given by $1 - \sigma = \frac{P(\mathbf{x},\mathbf{y}) - W(\mathbf{x},\mathbf{y})}{P(\mathbf{x},\mathbf{y}) - U(\mathbf{x})}$.

Wages are set as follows. Wage contracts are renogotiated sequentially as in Postel-Vinay and Robin (2002). A worker accepts the outside offer if $P(\mathbf{x}, \mathbf{y}') > P(\mathbf{x}, \mathbf{y})$. In this case, their new wage contract is worth $W' = P(\mathbf{x}, \mathbf{y})$ as they are able to extract $\sigma' = \frac{P(\mathbf{x},\mathbf{y}) - U(\mathbf{x})}{P(\mathbf{x},\mathbf{y}') - U(\mathbf{x})}$ from the outside offer. On the other hand, a worker rejects the outside offer if $P(\mathbf{x}, \mathbf{y}') < P(\mathbf{x}, \mathbf{y})$. In this case, they can still receive a higher wage if the value of the new match is still greater than their current value (i.e., $P(\mathbf{x}, \mathbf{y}) \geq P(\mathbf{x}, \mathbf{y}') > W(\mathbf{x}, \mathbf{y})$). In this case, the worker renegotiates such that they stay in their current job, but receives a new wage contract equal to the match value of the outside offer (i.e., $W' = P(\mathbf{x}, \mathbf{y}')$). Under this negotiation, the worker receives a higher share of the original surplus, this time given by $\sigma' = \frac{P(\mathbf{x},\mathbf{y}')-U(\mathbf{x})}{P(\mathbf{x},\mathbf{y})-U(\mathbf{x})}$.

As discussed in Lise and Postel-Vinay (2020), the increase in surplus can be given to the worker in various ways. Following Lise and Postel-Vinay (2020), I let the worker's skills and current match surplus evolve in their current match per usual, but use the new share σ' to determine the worker's wage until a new job negotiation or job transition occurs. This set-up implies that workers only receive rent through Bertrand competition between employers, as in Lise and Postel-Vinay (2020) and originally Postel-Vinay and Robin (2002). This set-up results in workers having low wages when first entering employment from unemployment and then potentially large wage growth as they receive outside offers.

When receiving a job offer out of unemployment, the worker only extracts the value of unemployment from the offer (assuming the value of the new job is greater), and so thus $\sigma = 0$. When already employed, the worker's surplus remains 0 unless they receive an outside job offer, in which case the surplus share is given by the following, as discussed in the previous paragraph:

$$\sigma_t = \begin{cases} \frac{P(\mathbf{x},\mathbf{y})-U(\mathbf{x})}{P(\mathbf{x},\mathbf{y}')-U(\mathbf{x})} & \text{if } P(\mathbf{x},\mathbf{y}') > P(\mathbf{x},\mathbf{y}) \\ \frac{P(\mathbf{x},\mathbf{y}')-U(\mathbf{x})}{P(\mathbf{x},\mathbf{y})-U(\mathbf{x})} & \text{if } P(\mathbf{x},\mathbf{y}) \geq P(\mathbf{x},\mathbf{y}') > W(\mathbf{x},\mathbf{y}) \\ \sigma_{t-1} & \text{if } P(\mathbf{x},\mathbf{y}) > W(\mathbf{x},\mathbf{y}) > P(\mathbf{x},\mathbf{y}') \end{cases} . \tag{21}$$

The total match value depends on the value of output, the disutility of work, the value of unemployment, and the gains from skill accumulation. In particular, it is given by

$$\begin{aligned} (r+\delta)P(\mathbf{x},\mathbf{y}) &= f(\mathbf{x},\mathbf{y},p) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x}) + h(\mathbf{x},\mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x},\mathbf{y}) \\ &\quad + \lambda_i^1(i) E_{\mathbf{y}'\sim\Upsilon}[W(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}) | P(\mathbf{x},\mathbf{y}') > P(\mathbf{x},\mathbf{y})] \\ &= f(\mathbf{x},\mathbf{y},p) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x}) + h(\mathbf{x},\mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x},\mathbf{y}). \end{aligned} \tag{22}$$

Note that from here on I denote $\delta(\mathbf{x}, \mathbf{y}, p)$ by δ for sake of exposition. The term $E_{\mathbf{y}'\sim\Upsilon}[W(\mathbf{x},\mathbf{y}')-P(\mathbf{x},\mathbf{y})|P(\mathbf{x},\mathbf{y}') > P(\mathbf{x},\mathbf{y})]$ equals zero because of the following. If the individual receives an outside offer and stays in the current job, then the value of the new wage $W(\mathbf{x},\mathbf{y}')$ is $P(\mathbf{x},\mathbf{y})$, in which case $W(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}) = 0$. On the other hand, if they accept the outside offer, they receive the same value $P(\mathbf{x},\mathbf{y})$, but from the new job. In this case as well, $W(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}) = 0$. Thus, in either case, $W(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}) = 0$ and so therefore the second line of equation 22 equals zero.

The value for an employed worker is the sum of the value of unemployment plus the share of surplus the worker receives. In particular, it is given by

$$\begin{aligned} W(\mathbf{x}, \mathbf{y}, \sigma) &= U(\mathbf{x}) + \sigma(P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x})) \\ &= (1-\sigma)U(\mathbf{x}) + \sigma P(\mathbf{x}, \mathbf{y}). \end{aligned} \tag{23}$$

The worker value function is given by:

$$\begin{aligned} (r+\delta)W(\mathbf{x}, \mathbf{y}, \sigma) = w(\mathbf{x}, \mathbf{y}, \sigma) - c(\mathbf{x}, \mathbf{y}, p) + \delta U(\mathbf{x}) \\ + \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x}, \mathbf{y}'), P(\mathbf{x}, \mathbf{y})\} - W(\mathbf{x}, \mathbf{y}, \sigma)\}] \\ + h(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} W(\mathbf{x}, \mathbf{y}, \sigma). \end{aligned} \tag{24}$$

The value for an unemployed worker depends on the flow utility of unemployment as well as the extent to which skills change in unemployment. It is given by

$$rU(\mathbf{x}) = b(\mathbf{x}) + h(\mathbf{x}, \mathbf{0}) \cdot \nabla U(\mathbf{x}). \tag{25}$$

Combining equations 22, 23, 24, and 25 gives the following wage equation (see Appendix G.1 for the derivation):

$$\begin{aligned} w(\mathbf{x}, \mathbf{y}, \sigma) = \sigma f(\mathbf{x}, \mathbf{y}, p) + (1-\sigma)(b(\mathbf{x}) + c(\mathbf{x}, \mathbf{y}, p)) \\ - \lambda_1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x}, \mathbf{y}') - P(\mathbf{x}, \mathbf{y}), 0\} + (1-\sigma(P(\mathbf{x}, \mathbf{y}) - U(\mathbf{x}))\}] \\ - (1-\sigma)(h(\mathbf{x}, \mathbf{y}) - h(\mathbf{x}, \mathbf{0})) \cdot \nabla U(\mathbf{x}). \end{aligned} \tag{26}$$

This equation is identical to that in Lise and Postel-Vinay (2020), except in the dependence on background characteristics p, as well as how the occupational sampling distribution Υ is defined. Importantly, letting the background characteristics p impact both production and disutility has simple and important implications for wages and what jobs are viewed as desirable and acceptable to workers.

Note that the wage depends positively on match output $f(\mathbf{x}, \mathbf{y}, p)$ as well as home production $b(\mathbf{x})$ and the disutility of work $c(\mathbf{x}, \mathbf{y}, p)$, as workers are compensated for their foregone respective values. This implies that if workers have an idiosyncratic preference for their parent's occupation (such that their disutility from their parent's occupation is lower), they will be willing to take a lower wage for the same occupation. Lastly, wages depend negatively on the value of future outside offers and on the value of skill accumulation.

G Model derivations

As discussed in the main text, the labor market stage of the model draws heavily from Lise and Postel-Vinay (2020). As a result, the model derivations below also draw heavily from Lise and Postel-Vinay (2020).

G.1 Wage equation

I start with the worker value function from equation 24,

$$\begin{aligned}(r+\delta)W(\mathbf{x},\mathbf{y},\sigma) = w(\mathbf{x},\mathbf{y},\sigma) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x})\\ + \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0,\min\{P(\mathbf{x},\mathbf{y}'),P(\mathbf{x},\mathbf{y})\} - W(\mathbf{x},\mathbf{y},\sigma)\}]\\ + h(\mathbf{x},\mathbf{y})\cdot\nabla_{\mathbf{x}}W(\mathbf{x},\mathbf{y},\sigma),\end{aligned}$$

and substitute equation 23 for $W(\mathbf{x},\mathbf{y},\sigma)$:

$$\begin{aligned}(r+\delta)((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y})) = w(\mathbf{x},\mathbf{y},\sigma) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x})\\ + \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0,\min\{P(\mathbf{x},\mathbf{y}'),P(\mathbf{x},\mathbf{y})\}\\ - ((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y}))\}]\\ + h(\mathbf{x},\mathbf{y})\cdot\nabla_{\mathbf{x}}((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y})).\end{aligned}$$

This can be re-written as

$$\begin{aligned}w(\mathbf{x},\mathbf{y},\sigma) = (r+\delta)((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y})) + c(\mathbf{x},\mathbf{y},p) - \delta U(\mathbf{x})\\ - \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0,\min\{P(\mathbf{x},\mathbf{y}'),P(\mathbf{x},\mathbf{y})\} - ((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y}))\}]\\ - h(\mathbf{x},\mathbf{y})\cdot\nabla_{\mathbf{x}}((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y})),\end{aligned}$$

which can be further simplified and re-written as

$$\begin{aligned}w(\mathbf{x},\mathbf{y},\sigma) = (r+\delta)(1-\sigma)U(\mathbf{x}) + c(\mathbf{x},\mathbf{y},p) - \delta U(\mathbf{x})\\ - \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0,\min\{P(\mathbf{x},\mathbf{y}')-P(\mathbf{x},\mathbf{y}),0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y})-U(\mathbf{x}))\}]\\ - h(\mathbf{x},\mathbf{y})\cdot\nabla_{\mathbf{x}}((1-\sigma)U(\mathbf{x})+\sigma P(\mathbf{x},\mathbf{y})) + \sigma(r+\delta)P(\mathbf{x},\mathbf{y}).\end{aligned}$$

I then substitute equation 22 for $(r+\delta)P(\mathbf{x},\mathbf{y})$ to get

$$\begin{aligned}
w(\mathbf{x},\mathbf{y},\sigma) = (r+\delta)(1-\sigma)U(\mathbf{x}) + c(\mathbf{x},\mathbf{y},p) - \delta U(\mathbf{x})& \\
- \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}), 0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y}) - U(\mathbf{x}))\}]& \\
- h(\mathbf{x},\mathbf{y}) \cdot \nabla_{\mathbf{x}}((1-\sigma)U(\mathbf{x}) + \sigma P(\mathbf{x},\mathbf{y}))& \\
\sigma(f(\mathbf{x},\mathbf{y},p) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x}) + h(\mathbf{x},\mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x},\mathbf{y})).&
\end{aligned}$$

This simplifies to

$$\begin{aligned}
w(\mathbf{x},\mathbf{y},\sigma) = (r+\delta)(1-\sigma)U(\mathbf{x}) + c(\mathbf{x},\mathbf{y},p) - \delta U(\mathbf{x})& \\
- \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}), 0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y}) - U(\mathbf{x}))\}]& \\
- h(\mathbf{x},\mathbf{y}) \cdot \nabla_{\mathbf{x}}((1-\sigma)U(\mathbf{x}))& \\
\sigma(f(\mathbf{x},\mathbf{y},p) - c(\mathbf{x},\mathbf{y},p) + \delta U(\mathbf{x})),&
\end{aligned}$$

and can then be re-written as (after also removing the subscript on $\nabla_{\mathbf{x}}$)

$$\begin{aligned}
w(\mathbf{x},\mathbf{y},\sigma) = \sigma f(\mathbf{x},\mathbf{y},p) + (1-\sigma)c(\mathbf{x},\mathbf{y},p) + (1-\sigma)rU(\mathbf{x})& \\
- \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}), 0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y}) - U(\mathbf{x}))\}]& \\
- h(\mathbf{x},\mathbf{y})(1-\sigma) \cdot \nabla U(\mathbf{x}).&
\end{aligned}$$

I can then substitute in equation 25 for $rU(x)$

$$\begin{aligned}
w(\mathbf{x},\mathbf{y},\sigma) =& \sigma f(\mathbf{x},\mathbf{y},p) + (1-\sigma)c(\mathbf{x},\mathbf{y},p) + (1-\sigma)b(\mathbf{x}) \\
&- \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}), 0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y}) - U(\mathbf{x}))\}] \\
&- (1-\sigma)(h(\mathbf{x},\mathbf{y}) - h(\mathbf{x},\mathbf{0})) \cdot \nabla U(\mathbf{x}),
\end{aligned}$$

which can then further be simplified to

$$\begin{aligned}
w(\mathbf{x},\mathbf{y},\sigma) =& \sigma f(\mathbf{x},\mathbf{y},p) + (1-\sigma)(b(\mathbf{x}) + c(\mathbf{x},\mathbf{y},p)) \\
&- \lambda_i^1 E_{\mathbf{y}'\sim\Upsilon}[\max\{0, \min\{P(\mathbf{x},\mathbf{y}') - P(\mathbf{x},\mathbf{y}), 0\} + (1-\sigma)(P(\mathbf{x},\mathbf{y}) - U(\mathbf{x}))\}] \\
&- (1-\sigma)(h(\mathbf{x},\mathbf{y}) - h(\mathbf{x},\mathbf{0})) \cdot \nabla U(\mathbf{x}).
\end{aligned}$$

This is the final wage equation, which is shown as equation 26 in Appendix Section F.

G.2 Skill accumulation

I begin with a generic skill dimension k and solve $\dot{x}_k = \gamma_k(y_k - x_k)$. This can be written as $\frac{dx_k}{dt} = \gamma_k(y_k - x_k)$ or $\frac{1}{y_k - x_k}dx_k = \gamma_k dt$. Integrating both sides gives $-\ln|x_k - y_k| = \gamma_k t + C$, where C is the constant of integration. I then solve this for x. It can first be written as $e^{-\gamma_k t - C} = |y_k - x_k|$, which then gives $x_k(t) = y_k - e^{-\gamma_k t - C}$(assuming $y > x$, without loss of generality). Given that C is just a constant, this can be written as $x_k(t) = y_k - Ce^{-\gamma_k t}$. I now rewrite this in order to make it more tractable for the estimation. I first re-write this at time $t = s$ as $C = (y_k - x_k(s))e^{\gamma_k s}$. I then substitute this back into the initial equation $x_k(t) = y_k - Ce^{-\gamma_k t}$ to get $x_k(t) = y_k - (y_k - x_k(s))e^{\gamma_k s})e^{-\gamma_k t}$, which can be simplified to $x_k(t) = y_k - e^{-\gamma_k(t-s)}(y_k - x_k(s))$, which is the final specification. Skill accumulation over one period, from t to $t+1$, can then be written as $x(t+1) = y - e^{-\gamma_k}(y - x(t))$.

I now solve for the general efficiency of the worker. I start with the differential equation $\dot{x_T} = gx_T$. I guess the solution $x_T(t) = x_T(0)e^{gt}$, which gives $\dot{x_T} = x_T(0)ge^{gt} = gx_T(0)e^{gt} = gx_T(t)$, which is indeed correct.

G.3 Value of unemployment

We can guess and verify the solution to equation 25 for the value of unemployment. Note that the value of unemployment $U(x) = \frac{bx_T}{r-g}$ is independent of the multidimensional skills $\mathbf{x} = (x_q, x_v, x_s, x_m)$ and so $\nabla U(\mathbf{x}) = (0, 0, 0, 0, \frac{b}{r-g})$. I then verify equation 25 as $r \cdot \frac{bx_T}{r-g} = bx_T + \begin{pmatrix} . \\ . \\ . \\ . \\ gx_T \end{pmatrix}^T \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{b}{r-g} \end{pmatrix}$, which can be simplified to $\frac{rbx_T}{r-g} = bx_T + gx_T\frac{b}{r-g}$, which is correct, thus verifying the guess.

G.4 Total match surplus

In this section, I solve for the total match surplus, largely following the proof from Lise and Postel-Vinay (2020), with the relevant modifications. Starting with the equation that defines the match value (equation 22), we have

$$(r + \delta)P(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y}, p) - c(\mathbf{x}, \mathbf{y}, p) + \delta U(\mathbf{x}) + h(\mathbf{x}, \mathbf{y}) \cdot \nabla_{\mathbf{x}} P(\mathbf{x}, \mathbf{y}).$$

This is a first order linear partial differential equation. The solution is characterized by the following system of $K+1$ ordinary differential equations

$$\begin{aligned}\frac{dx_k}{dt} =& h_k(\mathbf{x}(t), \mathbf{y}), k = 1\ldots, K \\ \frac{dw}{dt} =& (r+\delta)w - [f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p)] - \delta U(\mathbf{x}(t)),\end{aligned}$$

where $h_k(\mathbf{x}(t), \mathbf{y})$ is given by equation 8 in the main text (Section 4.2).

The total match value is then given by $P(\mathbf{x}(t), \mathbf{y}) = w(t)$. Initial conditions are given by the initial skills of workers $\mathbf{x}(0)$. I denote the solution to the first equation as $\mathbf{X}(t; \mathbf{y}, \mathbf{x}_0)$, given initial conditions $\mathbf{x}_0$ and job type $\mathbf{y}$. The second equation can be rearranged to

$$(r+\delta)w - \frac{dw}{dt} = f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t)).$$

Multiplying both sides by $e^{-(r+\delta)t}$ gives

$$e^{-(r+\delta)t}((r+\delta)w - \frac{dw}{dt}) = e^{-(r+\delta)t}(f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t))),$$

of which the left-hand side can then be re-written as

$$-\frac{d}{dt}(e^{-(r+\delta)t}w) = e^{-(r+\delta)t}(f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t))).$$

Integrating both sides then gives

$$-\int_t^\infty \frac{d}{dt}(e^{-(r+\delta)t}w)ds = \int_t^\infty e^{-(r+\delta)t}(f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t)))ds.$$

This can then be simplified to

$$e^{-(r+\delta)t}w = \int_t^\infty e^{-(r+\delta)s}(f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t)))ds.$$

Finally, this can be written as

$$P(\mathbf{x}(t), \mathbf{y}) = w = \int_t^\infty e^{-(r+\delta)(s-t)}(f(\mathbf{x}(t), \mathbf{y}, p) - c(\mathbf{x}(t), y, p) + \delta U(\mathbf{x}(t)))ds.$$

Using the solution from the first partial differential equation $\mathbf{X}(t; \mathbf{y}, \mathbf{x}_0)$, the solution to the second partial differential equation is then given by

$$P(\mathbf{x}(t),\mathbf{y}) = \int_t^\infty [f(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t)),\mathbf{y},p) - c(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t)),\mathbf{y},p) + \delta U(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t)))]e^{-(r+\delta)(s-t)}ds.$$

The value of unemployment can be solved similarly (Lise and Postel-Vinay (2020)), to get

$$U(\mathbf{x}(t)) = \int_t^\infty b(\mathbf{X}(s;0,\mathbf{x}(t)))e^{-r(s-t)}ds.$$

The difference, or total surplus, is then given by

$$P(\mathbf{x}(t),\mathbf{y}) - U(\mathbf{x}(t)) = \int_t^\infty [f(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t)),\mathbf{y},p) - c(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t)),\mathbf{y},p) - b(\mathbf{X}(s;\mathbf{y},\mathbf{x}(t))]e^{-(r+\delta)(s-t)}ds.$$

I now substitute in the functional form assumptions for the production function and disutility function from Section 4.2 to get

$$\begin{aligned} P(\mathbf{x}(t),\mathbf{y}) - U(\mathbf{x}(t)) = \int_t^\infty &[x_T(t) \times (\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x x_k(t) + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}) \\ &- x_T(t) \times (\pi_2 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}) \\ &- b x_T(t)]e^{-(r+\delta)(s-t)}ds. \end{aligned}$$

This can be written as

$$\begin{aligned} P(\mathbf{x}(t),\mathbf{y}) - U(\mathbf{x}(t)) = \int_t^\infty &[\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x x_k(t) + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} \\ &- \pi_2 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} \\ &- b] x_T(t) e^{-(r+\delta)(s-t)}ds. \end{aligned}$$

We can then substitute to convert the index from t to s using the formula for the general efficiency of the worker (in Section 4.2), which then gives

$$\begin{aligned} P(\mathbf{x}(t),\mathbf{y}) - U(\mathbf{x}(t)) = \int_t^\infty &[\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x x_k(t) + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} \\ &- \pi_2 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} \\ &- b] x_T(s) e^{-g(s-t)} e^{-(r+\delta)(s-t)}ds. \end{aligned}$$

We can then rewrite $x_k(t)$ in the first line as $x_k(t) = y_k - Ce^{-\gamma_k t}$ (see Appendix Section G.2) to get

$$P(\mathbf{x}(t), \mathbf{y}) - U(\mathbf{x}(t)) = \int_t^\infty [\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x (y_k - Ce^{-\gamma_k t}) + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}$$
$$- \pi_2 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}$$
$$- b] x_T(s) e^{-g(s-t)} e^{-(r+\delta)(s-t)} ds.$$

Finally, the solution is then given as:

$$P(\mathbf{x}(t), \mathbf{y}) - U(\mathbf{x}(t)) = x_T(t) \times \{ \frac{\alpha_T + \sum_{k=q,v,s,m} (\alpha_k^x y_k + \alpha_k^y y_k) + \pi_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} - \pi_2 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p} - b}{r + \delta - g}$$
$$+ \sum_{k=q,v,s,m} (\alpha_k^x) \times (\frac{\min\{x_k - y_k, 0\}}{r + \delta - g + \gamma_k} + \frac{\max\{x_k - y_k, 0\}}{r + \delta_i - g + \gamma_k}) \}.$$

This is the final expression for the surplus of a job match.

H Model estimation details

H.1 Pre-estimated parameters

Recall that in the first step of the estimation, I pre-estimate the parameters governing intergenerational multidimensional skill transmission, the occupational sampling distribution, and the job separation rate. In particular, I pre-estimate them as follows:

- Intergenerational multidimensional skill transmission: given parental occupation j, children draw from a skill k-specific normal distribution with mean and standard deviation taken from the actual estimation sample: $x_k \sim \mathcal{N}(\mu_{kj}, \sigma_{kj}^2)$ for $k = q, v, s, m$. For example, quantitative skills are drawn from $x_q \sim \mathcal{N}(\mu_{qj}, \sigma_{qj}^2)$, where μ_{qj} is the mean of quantitative skills for all children whose parents are in occupation j, while σ_{qj} is the standard deviation.

- Occupational sampling distribution: this is also taken from the estimation sample. In particular, I take directly from the data the likelihood of being in each two-digit occupation given one's college major.

- Job separation rate δ: this is taken as the mean from the estimation sample. Recall that the job separation rate is specified as $\delta(\mathbf{x}, \mathbf{y}, p) = \delta_0 + \delta_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}$ in the model specification. I take these values from the data, and so $\delta_0 = 0.0282$ and $\delta_1 = -0.0020$.

Lastly, I pre-set the annual discount rate r in the model as 0.10, following Lise and Postel-Vinay (2020).

H.2 More estimation details

H.2.1 Targeted moments

Below are the moments I target directly in the indirect inference.

Education stage The education stage moments are:

1. Shares in college, amongst those with father in occupation that does require college degree and does not require a college degree (defined as whether a majority in the occupation have a college degree or not)

2. Shares in college, amongst those that are in top half of skill distribution and bottom half (defined as simple average across all four skills)

3. Share with major with modal occupation same as parent's occupation

Labor market stage The labor market stage moments are:

1. Job transition rates: mean U2E rate and mean E2E rate across the entire sample

2. Mean wages across the life cycle: mean log wage at ages 24, 32, and 40

3. Wage regression: coefficients from a regression of log wages on age, tenure, worker skills, and occupational skill requirements

4. Worker skill and occupational skill requirement correlations: the correlation between worker skills and their occupational skill requirements for each skill $k = q, v, s, m$ at ages 32 and 40

5. Occupational following rate: the mean occupational following rate across the entire sample

6. Occupational follower wage premium: the difference in mean wages between followers and non-followers in the second time period[51]

[51] I use the second time period rather than the first time period because of the model structure. In particular, recall how wages are set in the model. Given that workers have no bargaining power when unemployed, workers' initial wages in a job spell tend to be low; but then they grow as they receive more job offers (see Lise and Postel-Vinay (2020) for more discussion). For this reason, I consider wages in the second time period rather than the first time period to compute the occupational following wage premium.

H.2.2 Weights

I follow Guvenen et al. (2021), among others, in how I construct weights. As in Guvenen et al. (2021), I do not use the optimal weighting matrix because the data set is very large and thus all the moments are very precisely estimated. Instead, I use the identity matrix, and then choose weights such that each set of moments is equally weighted. In particular, I use weights such that each grouping of moments as shown in Appendix Section H.2.1 are equally weighted.

H.2.3 Optimization routine

The optimization routine I use is the following. First, I use a global optimization algorithm. In particular, I use the controlled random search (CRS) with local mutation from the NLopt package. Second, given the solution to the global optimization, I use a local optimization algorithm (Nelder-Mead). Both algorithms are derivative-free.

H.2.4 Computing the standard errors

I compute asymptotic standard errors, following standard asymptotic theory (e.g., Duffie and Singleton (1993); Gourieroux et al. (1993); Pakes and Pollard (1989)). First, the variance-covariance matrix of $\hat{\theta}$ is given by

$$\Omega_\theta = (G'_\theta W G_\theta)^{-1} G'_\theta W [\Omega_g + \frac{N_d}{N_s}\Omega_g + G_\chi \Omega_\chi G'_\chi] W G_\theta (G'_\theta W G_\theta)^{-1},$$

where G_θ and G_χ are the gradient matrices of the moment conditions with respect to θ and χ, respectively, where again θ are the second-stage parameters and χ are the first-stage parameters; W is the weighting matrix; Ω_g is the variance-covariance matrix of the second-stage moment conditions; Ω_χ is the variance-covariance matrix of the first-stage parameter estimates; and N_d and N_s are the empirical sample size and the simulation sample size, respectively. The standard errors of the second-stage parameter estimates $\hat{\theta}$ are given by the square roots of the diagonal entries of Ω_θ.

I approximate the derivatives in the gradient matrix G_θ numerically. In particular, G_θ is the Jacobian matrix of the moment conditions with respect to θ, and is thus an M (number of moment conditions) by N (number of parameters) matrix. I approximate it with numerical methods, using a 2.5% step size with respect to the parameter estimates. I estimate Ω_g directly from the data, using 1,000 bootstraps of the original estimation sample data set. I use the weighting matrix W as described in Appendix Section H.2.2. Finally, I treat the first-stage parameters estimates as if they were known with certainty, thus letting $G_\chi = 0$.

These substitutions lead to the following

$$\Omega_\theta = (G'_\theta W G_\theta)^{-1} G'_\theta W [\Omega_g + \frac{N_d}{N_s}\Omega_g] W G_\theta (G'_\theta W G_\theta)^{-1},$$

which can then further be rewritten as below, which is the final expression for the variance-covariance matrix of $\hat{\theta}$:

$$\Omega_\theta = (1 + \frac{N_d}{N_s})(G'_\theta W G_\theta)^{-1} G'_\theta W \Omega_g W G_\theta (G'_\theta W G_\theta)^{-1}.$$

As mentioned above, the standard errors of the second-stage parameter estimates $\hat{\theta}$ are given by the square roots of the diagonal entries of Ω_θ.

H.3 Simulation details

I now discuss how I simulate the model. I simulate the model by creating a cohort of N workers over T years. Workers begin with an initial endowment of skills $\mathbf{x}$, as described above, which depend on their parent's occupation, p. Workers then make a college major choice, thus introducing a third main state variable, m, in addition to the first two (set of skills $\mathbf{x}$ and parental occupation p).

Workers next enter the labor market, in which each period their skills adjust accordingly to $\mathbf{h}(\mathbf{x}, \mathbf{y})$. Each worker is also then randomly faced with a labor market shock. If employed, they lose their job with probability δ and start the next period in unemployment, or they receive an outside offer with probability λ_1 and draw an outside offer with skill requirements $\mathbf{y}'$ from the sampling distribution $\Upsilon(\mathbf{y}|m)$. They decide to take the new job if the value of the outside offer exceeds the value of the current job (and thus are now in occupation $\mathbf{y}'$) or otherwise stay in their current job, with or without renegotiating the contract to match the outside offer. If the worker is unemployed, they also draw a job offer with probability λ_0 with skill requirements $\mathbf{y}'$ from the same sampling distribution $\Upsilon(\mathbf{y}|m)$; they accept the job offer if the value exceeds the value of unemployment. Lastly, the worker's wage is updated according to the wage equation 26 and model specification in Section 4.2.

Following Lise and Postel-Vinay (2020), I simulate the model over a presampling period in order to set the initial conditions. In the presampling period, all workers start as unemployed and then experience the same simulation as just described, except without skill accumulation or job separations. The presampling period ends when 63.2% of workers are employed, which is the employment rate in the estimation sample upon labor market entry. That point is then taken as the starting point in the actual simulation.

The output of the simulation is an $N \times T$ balanced panel of worker data that has the same format as the actual estimation sample data. The panel records the worker's skills, parental occupation, chosen college major, employment status and occupation, labor market transitions, and wages.

Other details I consider the first 17 years of the life cycle, due to data limitations. Note that Lise and Postel-Vinay (2020) use the first 15 years. Unlike in Lise and Postel-Vinay (2020), I abstract away from sample attrition for simplicity. I use as the data estimation sample the children born between 1975 and 1976, for a total of about 68,000 children (recall that I am limited to males as the Defence Force data is limited to males). In the simulation, I consider a sample of size 68,000 to mirror the data sample. I use the broad field of study categorization, as described in Appendix Section C, in the estimation. I do this for computational tractability, as considering the finer field of study categorization would take the estimation considerably longer and be infeasible.

H.4 Random sorting benchmark

The random sorting benchmark is calculated as follows. Consider the distribution of parent occupations, where the share of parents in occupation $j \in \mathcal{J}$ is given by p_j and such that $\sum_{j \in \mathcal{J}} p_j = 1$. If children were randomly assigned into occupations, then for the parents in occupation j, the likelihood they would have a child in their occupation would be the sample frequency of children in occupation j, which I denote s_j. The overall rate of occupational following then is a weighted mean of this share across all parent occupations, equal to $\sum_{j \in \mathcal{J}} p_j \cdot s_j$. At the two-digit level, the value of this in the model estimation sample is 4.5%. The value is 15.5% at the one-digit level and 50.7% at the class level.

H.5 Model counterfactuals

In Appendix Table A30 below, I discuss the way in which I implement each model counterfactual.

Table A30: Description of model counterfactuals

Mechanism	Change in model
Panel A. Education stage	
Equalize skill transmission	Instead of each child drawing from their parental occupation-specific distribution of skills (as described in Appendix Section H.1), each child now draws from the same distribution of skills. In particular, children now draw from a skill k-specific normal distribution with mean and standard deviation taken from the overall actual estimation sample: $x_k \sim \mathcal{N}(\mu_k, \sigma_k^2)$ for $k = q, v, s, m$. Note that now the distribution does not depend on parental occupation j, as it does in the baseline case
Equalize college cost	Set $\phi_1^c = 0$ from the general psychic cost of college (Equation 5)
No major cost	Set $\phi_1^m = 0$ from the major-specific psychic cost of college (Equation 6)
Panel B. Labor market stage	
No idiosyncratic preferences	Set $\pi_2 = 0$ in the specification of flow utility (Section 4.2)
No differential search frictions	In the main estimation, recall that the job separation rate is specified as $\delta(\mathbf{x}, \mathbf{y}, p) = \delta_0 + \delta_1 \mathbb{1}_{\mathbf{y}=\mathbf{y}_p}$, where $\delta_0 = 0.0282$ and $\delta_1 = -0.0020$. In this counterfactual, I specify $\delta(\mathbf{x}, \mathbf{y}, p) = \tilde{\delta_0}$, where $\tilde{\delta_0}$ is the overall job separation rate mean from the estimation sample and is equal to 0.0279
No occupation-specific skill	Set $\pi_1 = 0$ in the specification of production (Equation 7)

Notes: This table describes the precise manner in which each counterfactual is implemented.

H.6 Sensitivity analysis

I follow Andrews et al. (2017) and estimate the sensitivity of the parameter estimates $\hat{\theta}$ to the moment function $g(\hat{\theta})$. I estimate the sensitivity matrix as given by $\Lambda = (G'G)^{-1}G'$, which is an N (number of parameters) by M (number of moments) matrix and a mapping from the moments to the parameters. In Figures A10 and A11, I plot how a change in each moment would affect each parameter estimate. I group the moments into the groups defined in Section H.2.1. The results broadly align with the identification arguments in Section 5.2. The targeted moments are informative of the parameters they are meant to identify.

Figure A10: Andrews et al. (2017) sensitivity matrix for labor market parameters

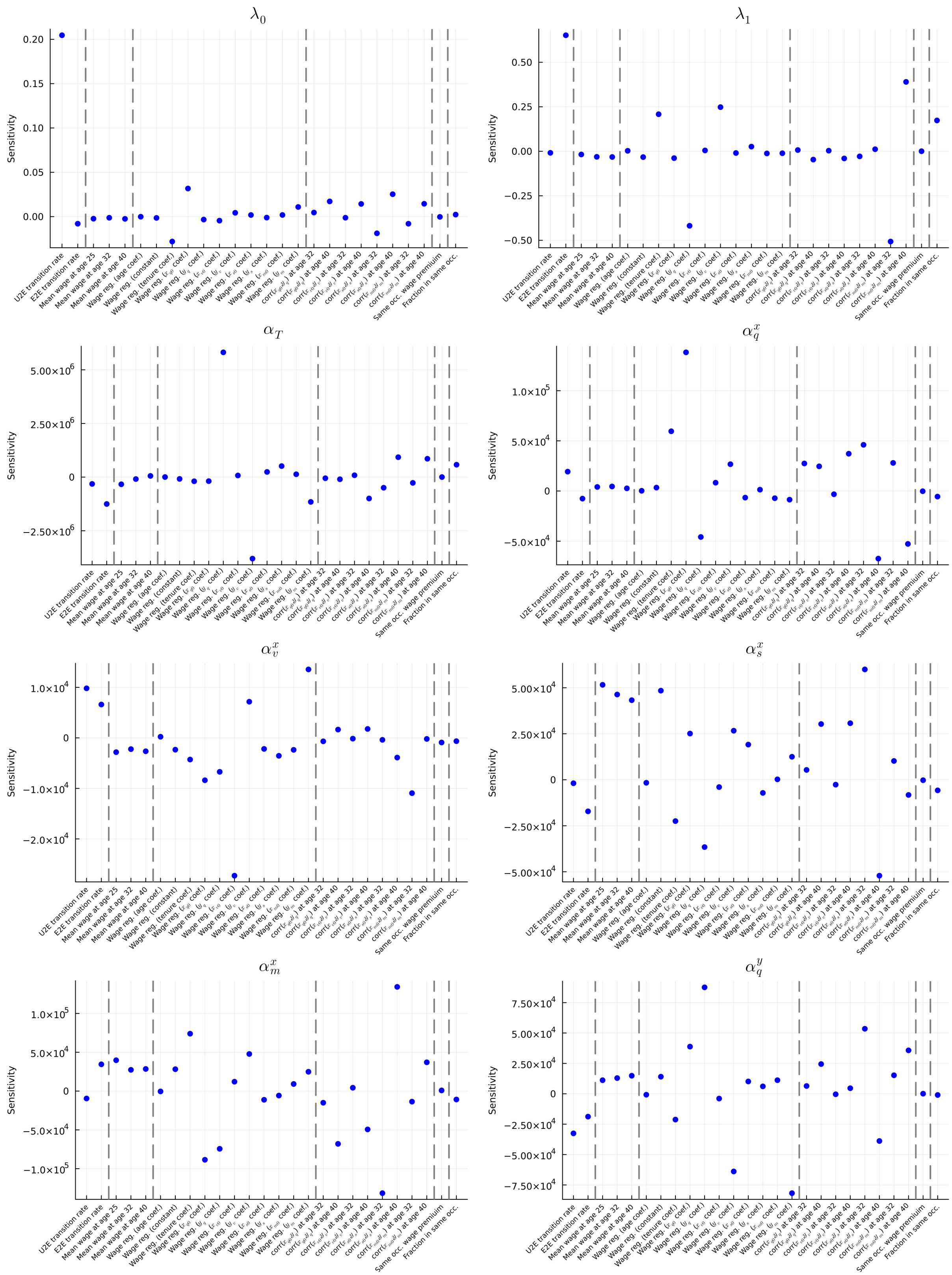


Figure A10: Andrews et al. (2017) sensitivity matrix for labor market parameters (cont.)

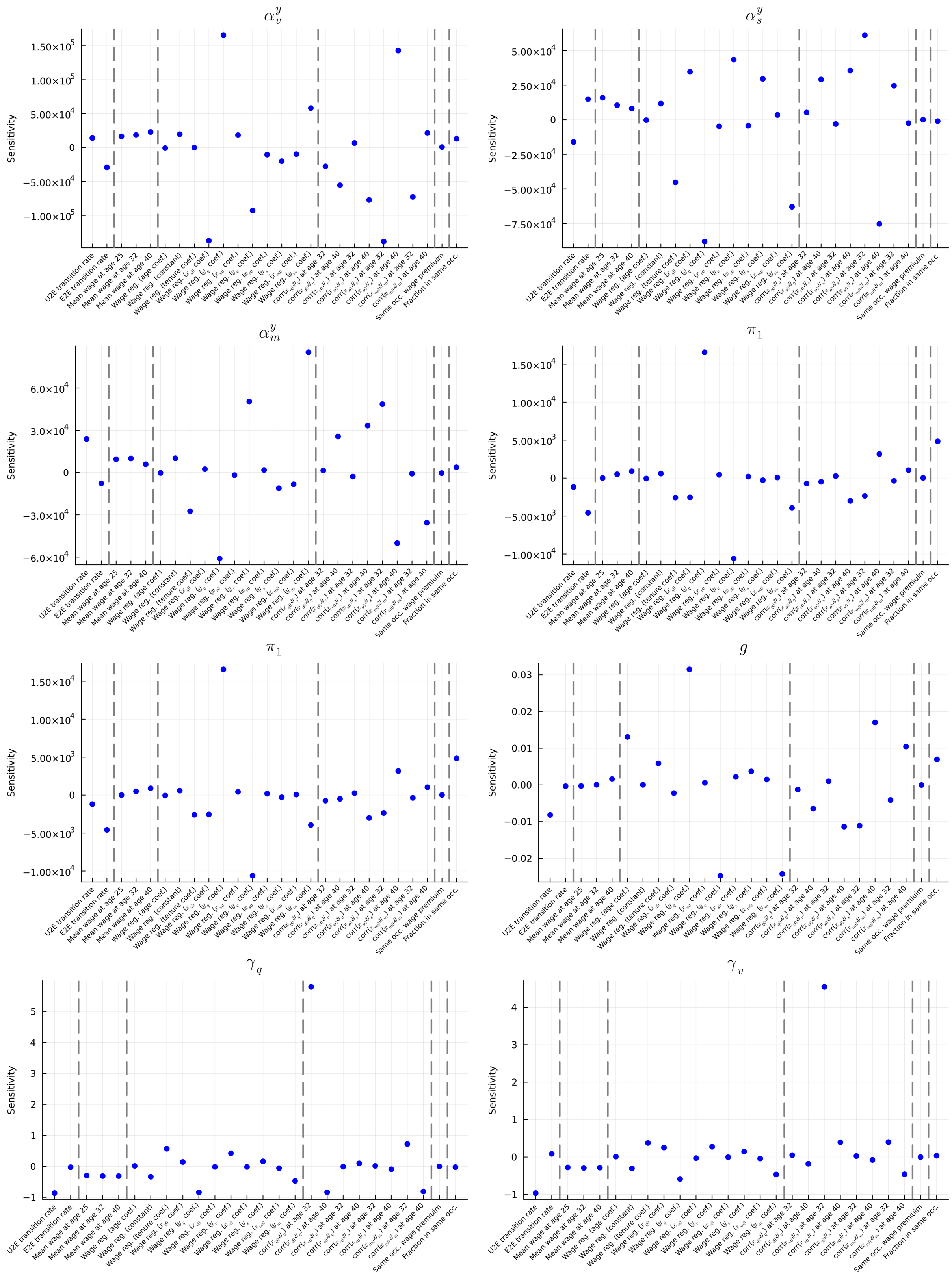
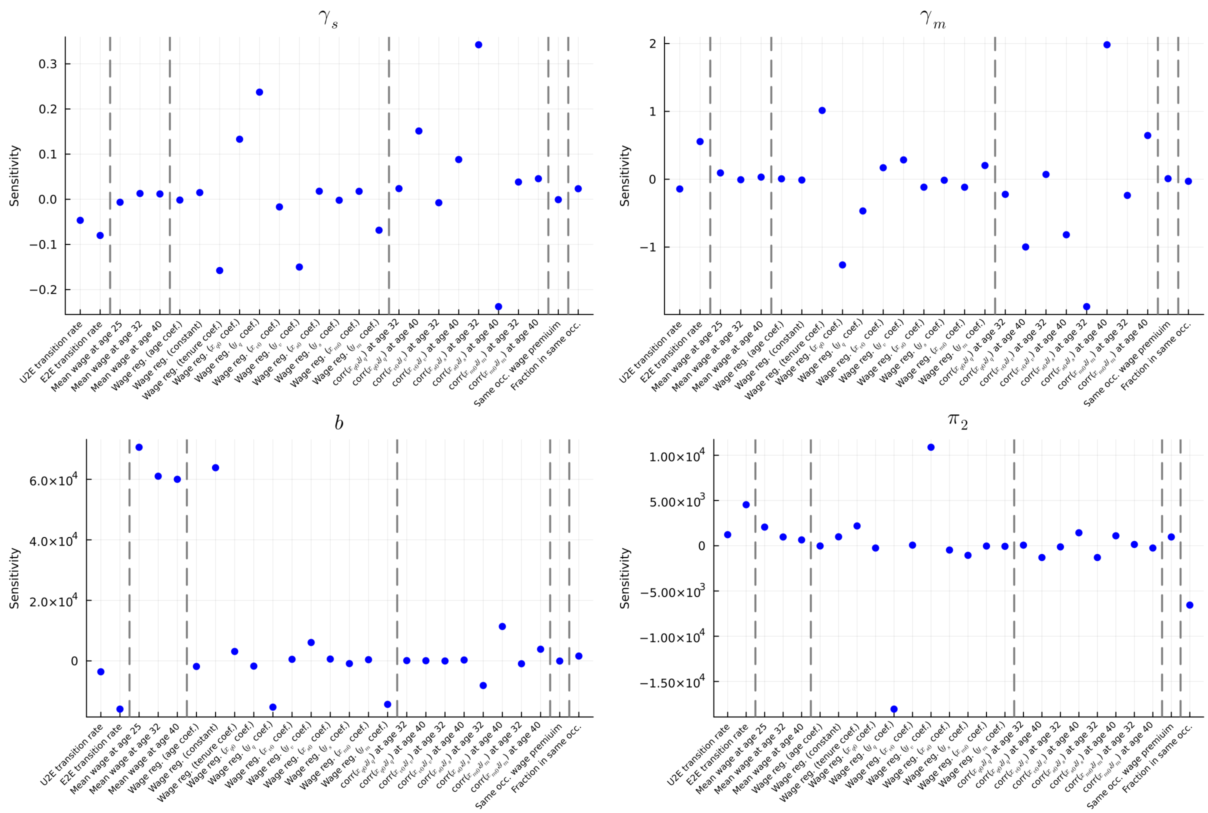
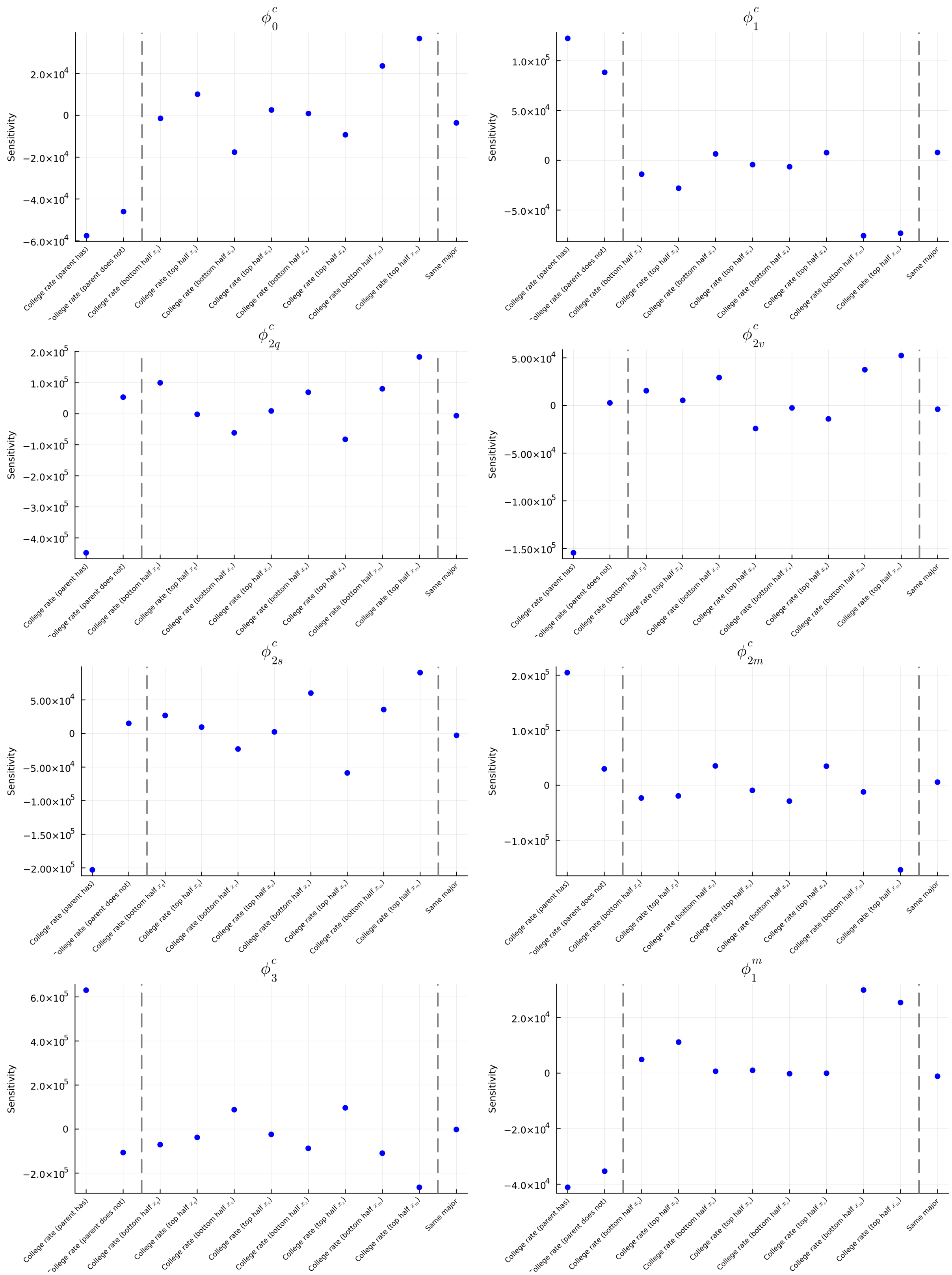


Figure A10: Andrews et al. (2017) sensitivity matrix for labor market parameters (cont.)



Notes: Figures illustrate results from Andrews et al. (2017) sensitivity matrix. Each figure is for a given parameter, with all moments on horizontal axis.

Figure A11: Andrews et al. (2017) sensitivity matrix for education parameters



Notes: Figures illustrate results from Andrews et al. (2017) sensitivity matrix. Each figure is for a given parameter, with all moments on horizontal axis.

I Robustness of model results

I.1 Example of general equilibrium responses

Table 10 from Abbott et al. (2019) considers three different potential expansions of federal student aid policies. The table shows how college attainment would change both under a short run partial equilibrium and a long run general equilibrium in which wages can respond to changes in labor supply and there are capacity constraints in labor demand. The table also shows how college attainment would differentially change for children of parents in the top-third of the wealth distribution and children of parents in the bottom-third of the wealth distribution. I include their estimates in Table A31 and then calculate the class gap in college attainment for each of these policies and in both partial and general equilibrium, and how the class gap changes relative to the baseline case. In Panel A, for example, the class gap decreases by 39% in partial equilibrium. When considering general equilibrium responses, it decreases by 16%, thus attenuating the partial equilibrium reduction by 58%. As shown in the other panels, the strength of the general equilibrium response varies across policies.

Table A31: General equilibrium responses in college completion

	Baseline	Partial Equilibrium	General Equilibrium
Panel A. General tuition grant expansion			
College attainment:			
Top third of parental wealth	0.399	0.416	0.405
Bottom third of parental wealth	0.205	0.263	0.226
Class gap:			
Gap	95%	58%	79%
Reduction relative to baseline		-39%	-16%
General equilibrium attenuation		-58%	
Panel B. Means-tested grant expansion			
College attainment:			
Top third of parental wealth	0.399	0.406	0.371
Bottom third of parental wealth	0.205	0.288	0.248
Class gap:			
Gap	95%	41%	50%
Reduction relative to baseline		-57%	-48%
General equilibrium attenuation		-16%	
Panel C. Merit-based grant expansion			
College attainment:			
Top third of parental wealth	0.399	0.412	0.405
Bottom third of parental wealth	0.205	0.249	0.225
Class gap:			
Gap	95%	65%	80%
Reduction relative to baseline		-31%	-15%
General equilibrium attenuation		-50%	

Notes: College attainment taken from Table 10 of Abbott et al. (2019). Class gap computed as percentage increase in likelihood of attaining college if in top third of parental wealth relative to being in bottom third. GE attenuation is percentage change in class gap reduction in GE relative to in PE.