Labor-market Drivers of Intergenerational Earnings Persistence

Erika Forsberg,[*] Martin Nybom[†] and Jan Stuhler[‡§]

August 21, 2024

Abstract

To what extent does the sorting of workers across firms contribute to intergenerational persistence, and why? We show that differences in firm pay premia account for 31% of the intergenerational elasticity of earnings in Sweden, rising to 38% when including dynamic returns to firm-specific experience. Firm pay gaps open already at career start, suggesting that children from more privileged backgrounds find more favorable entry points to the labor market. Their pay advantage widens further in early career as they climb the firm pay ladder faster, switch firms more frequently and secure higher pay gains conditional on switching. Skill sorting explains most of the widening in early career, but not the initial pay gaps at career start. These results are robust to accounting for compensating differentials and using alternative measures of firm quality.

[*]IFAU and Uppsala University (erika.forsberg@ifau.uu.se)

[†]IFAU (martin.nybom@ifau.uu.se)

[‡]Universidad Carlos III de Madrid (jan.stuhler@.uc3m.es)

[§]Financial support from Forte and the Ministerio de Ciencia e Innovación (Spain, RYC2019-027614-I) and Comunidad de Madrid (AMEPUC3M11 and H2019/HUM-5891) is gratefully acknowledged. We thank Olle Folke, Isaac Sorkin, Per Engzell, Guido Neidhöfer, Emran Shahe, Forhad Shilpi, Isabel Martínez, Rasmus Landersø, Jonas Cederlöf and seminar participants at IFAU, Uppsala University, and the Research Network for Intergenerational Mobility for helpful comments.

Introduction

Children from high-income families earn substantially higher incomes than children from low-income families (Jäntti and Jenkins 2014; Deutscher and Mazumder, 2023). A key question in the social sciences is why this is the case. Seminal contributions in the economic literature emphasize parental investments in human capital (Becker and Tomes, 1979), dynamic complementarities in skill development across childhood (Heckman and Mosso, 2014), and the role of credit constraints in these processes (Lochner and Monge-Naranjo, 2012; Lee and Seshadri, 2019). A related strand of the literature studies the roles of nature and nurture, decomposing intergenerational transmission into pre- and post-birth factors (e.g. Björklund et al., 2006; Black et al., 2005). As such, much of the existing literature focuses on how differences in human capital accumulation *in childhood* translate to pay differences in adulthood. In contrast, very few studies focus on the role of labor-market mechanisms in shaping intergenerational persistence in incomes.

In this paper, we study the contribution of firms, sorting, and career advancement on the labor market to intergenerational persistence. Employing population-wide earnings data with employer-employee and parent-child links from Sweden, we decompose earnings into individual and firm components using the two-way fixed effects "AKM" framework of Abowd et al. (1999). We then show that labor-market factors, broadly speaking, and socioeconomic status (SES) gradients in sorting across employers more specifically, explain 31% of overall income persistence measured at mid-age. The contribution of firms rises to 38% if we incorporate differences in returns to experience across firm types in our analysis. The importance of such labor-market mechanisms grows over age, suggesting that differences in labor-market sorting can partly explain why SES gradients in earnings increase with age. The SES gradient in firm pay (or the "firm pay gradient") can partly be attributed to confounding sorting on skills, but at least half of the gradient remains when conditioning on a rich set of skill measures. Finally, we analyze whether there is a similar gradient in terms of the overall attractiveness of employers, including their non-pay characteristics. We find little evidence that the purely monetary focus overstates intergenerational persistence in overall welfare.

We first present a set of empirical facts on the SES gradient in firm pay premia in Sweden, adopting a similar approach as an earlier study on Israeli data by Dobbin and Zohar (2023). Decomposing earnings into worker- and firm-specific components, we show that the "static" firm component accounts for nearly one third (31%) of the intergenerational earnings elasticity (IGE). This estimated contribution of firms to intergenerational persistence is in line with recent evidence by Engzell and Wilmers (2024) for Sweden, but substantially higher than the corresponding estimates for Israel by Dobbin and Zohar (2023). We provide evidence suggesting that the high quality of our data, spanning over a 35 year period, contributes to this difference.

The sorting of workers across firms therefore explains a substantial part of the transmission of income advantages from one generation to the next. Selection into higher-paying industries, such as the financial sector or IT services, explains some of the firm pay gap, while differences in firm pay by region or firm size are less important. Jointly, region, industry, and firm size explain more than half of the SES gradient in firm pay. Where one works plays a particularly important role for intergenerational persistence at the top of the parental income distribution. While both the worker and firm components increase with parental income, the firm component increases more strongly in relative terms. As such, this steep rise in firm premia in the right tail of the parent-income distribution is a potentially important explanation for the elevated persistence at the top of Sweden's income distribution, as documented by Björklund et al. (2012).

In the second part of our paper, we study how the firm pay gradient evolves over the lifecycle. Do children from high-income parents sort into higher-paying firms right at the beginning of their career, as one would expect if parents (Kramarz and Skans, 2014; Staiger, 2022) or peers (Cornelissen et al., 2017; Zimmerman, 2019) provide access to higher-paying firms? Or does the firm pay gap widen gradually, as one might expect if children from high-income families make faster career progress, perhaps by exploiting social and informational networks and by navigating the job search process more proficiently? We find that much of the disparities in firm pay open up right at the beginning of the career, consistent with the interpretation that children from richer parents find more favorable "entry points" to the labor market. But the SES gradient in firm pay then widens further in the early part of the career, before stabilizing in the mid-30s. Differences in early-career progress therefore explain part of the firm pay gap, coinciding with the career stage when many workers experience high earnings growth and when differences in earnings growth magnify relatively fast.

Indeed, high-SES children not only more frequently switch firms in the early career, but also make larger gains in firm pay conditional on switching. These results can be interpreted using the family of models that illustrate how search frictions can lead to variation in firm pay and "job ladders" (Burdett and Mortensen, 1998; Manning, 2013), suggesting that high-SES children reach higher rungs of the firm job ladder faster. All these patterns hold within education groups, but are especially strong among college graduates. These differences in the frequency and quality of job switching also hold when conditioning on voluntary firm switches, which we define as firm-to-firm transitions without any intermediate unemployment spell. Moreover, firm pay gradients exist not only in static comparisons ("firm fixed effects"), but also in a dynamic sense: children from more privileged backgrounds sort into expanding firms that are characterized by higher wage *growth*.

To study the implication of such dynamic firm advantages, we follow Arellano-Bover and Saltiel (2021) and use a k-means clustering algorithm to estimate an extended two-way fixed effects model that allows for firm-class specific returns to experience. We show that children from affluent families sort into firms with substantially higher wage growth; by age 40, these dynamic firm advantages accumulate to 20% of their overall firm pay advantage. Incorporating these dynamic advantages, the total contribution of firms increases to 35% of the IGE (or 38% when netting out the AKM residuals and covariates). We therefore find that labor market sorting explains a large share of the intergenerational persistence in income.

In the third part of our paper, we investigate whether the overrepresentation of high-SES children in better-paying firms is due to their higher skill levels. It is well established that firm and individual components from the AKM framework correlate positively due to assortative matching, where more productive workers sort into higher-paying firms (Card et al., 2013). Given that high-SES children have on average higher individual earnings components, the SES gradient in firm pay might simply reflect this skill-based sorting. In this case, the firm's role would primarily be that of a "mediator" for the effect of skill differences and inequities that already arose in childhood.

We first show that the SES gradient in firm pay weakens when conditioning on the worker fixed effects from the AKM model: a unit increase in log parental income increases the firm pay premium at age 40 by 3.7 rather than 5.4 log points in the unconditional case. Based on these results, about 30% of the SES gradient in firm pay can be attributed to skill sorting as captured by the worker fixed effects – a similar share as Dobbin and Zohar (2023) find for Israel, using the same approach. However, our main contribution here is that we can test for skill sorting more directly, as we have access to late-adolescence skill scores from the mandatory military enlistment tests. Conditioning on cognitive and non-cognitive skills measured at age 18, the residual firm pay gradient decreases to about 50% of the unconditional SES gradient, i.e. *half* of the SES gradient in firm pay at age 40 is due to skill sorting. The cognitive skill measure provides the most additional explanatory power to understand the pattern of sorting.

While this analysis confirms that a major share of the SES gradient in firm pay is due to skill sorting, we find that family background plays an important role for labor-market outcomes even conditional on skill. As is perhaps intuitive, family background plays a particularly important role in the early career: at age 25, nearly 70% of the firm pay gradient is due to direct family effects unrelated to skill, falling to just 50% at age 40. Parental background and networks thus provide particularly crucial advantages at labor market entry; this finding is in line with evidence by Kramarz and Skans (2014), San (2022) and Staiger (2022), who document the importance of parents' co-worker networks for their children's job finding. These studies focus on isolating the specific effects of parental networks on job finding and earnings among young workers. In comparison, we can quantify the overall contribution of family background on firm pay gradients, but cannot isolate the contribution of parental networks from other family background effects.

In the final part of our paper, we ask whether the remaining SES gradient in firm pay can be traced to differences in preferences, compensating differentials, and/or other non-wage attributes of firms. If intergenerational income transmission partly stems from inherited preferences – e.g. that some families value income and consumption compared to non-monetary attributes more than other families – then measures of income persistence would overstate the extent to which differences in welfare persist across generations. On the other hand, if children from more affluent families sort into jobs with more favorable non-monetary attributes, in addition to higher firm pay premia, then intergenerational mobility in underlying welfare would be even lower than income-based estimates suggest. To our knowledge, there exists very little evidence on this potential mechanism behind intergenerational income associations.

We test for the role of non-monetary attributes of firms, using different approaches. We first explore how firm premia and parental income relate to auxiliary measures of the attractiveness of firms, including quit rates and hiring rates from other firms (poaching). The results confirm that higher-paying firms are more desirable employers (in the sense of being able to poach workers from other firms), and that children from more affluent families sort into more desirable firms. To deepen this analysis, we employ a revealed-preferences based approach similar to Sorkin (2018). The approach exploits two-sided firm-to-firm transitions of workers over time, and infers the non-monetary values of firms from whether a firm on net gains or loses workers from firms of different values. Overall, we find little evidence on compensating differentials being systematically related to the family background of workers. Skills and other drivers of labor-market advantages seem thus more important for intergenerational transmission than correlated preferences for non-wage job attributes within families.

Our work contributes to a nascent literature on the importance of labor market factors to intergenerational persistence. An important motivation for our study is Dobbin and Zohar (2023), whose methodological approach we adopt in Section 2. Compared to their paper, we find a greater contribution of firms to intergenerational persistence; while the firm pay gradient contributes 22% to the IGE in their Israeli data, we find that firm advantages contribute 31% in our static and 38% in our dynamic model (at age 40, and net of AKM residuals and covariates). Differences in data quality and, in the dynamic model, the incorporation of firm-specific returns to experience explain most of this contrast. We further find that skill sorting explains half of the contribution of firm pay to the IGE. This estimate is in line with Dobbin and Zohar, even though we use a different approach to study skill sorting. Exploiting our rich data, we can extend their analysis in various ways. We study how the SES gradient in firm pay develops over the lifecycle, study the dynamic implications of labor market sorting, and examine the role of non-wage attributes of firms. Part of out analysis overlaps with Engzell and Wilmers (2024), who also study the intergenerational transmission of labor market advantages in Sweden. In comparison, they take a sociological perspective based on stratification theory and place more attention to the source of earnings in the parent generation.

A novel part of our paper is that we highlight the *dynamics* of firm pay gaps, and their development over the lifecycle. Our finding of a large SES gradient in firm pay opening upright at career start is consistent with studies tracing the effect of social connections on worker-firm sorting. In a pioneering study on the role of labor market factors in intergenerational transmission, Corak and Piraino (2010) document that many Canadian children work for the same employer as their parents, and that this association is stronger in high-income families; Bingley et al. (2012) and Stinson and Wignall (2018) confirm these findings for Denmark and the U.S. Further, Kramarz and Skans (2014) and Staiger (2022) quantify the role of parental and family connections in the early career. While their studies provide evidence on a particular mechanism (social connections), we quantify the contribution of firms to intergenerational persistence overall, which depends also on other mechanisms, such as skill sorting. We further show that rather than skills, the firm pay gradient at career start is primarily explained by other family-related advantages (including social connections). However, skill-based sorting explains most of the *increase* in the firm pay gradient over age.

Finally, our work relates to recent studies on skill sorting on the labor market (Eeckhout 2018, Card et al. 2018). Consistent with Dobbin and Zohar (2023), we find that there is far more sorting by parental background than one might expect from skill sorting alone. We further document firm sorting with respect to cognitive and non-cognitive skill measures (similar to Nybom, 2017, for college education), over and above the extent of sorting captured by worker effects from worker-firm two-way fixed effects models. Our results also relate to recent evidence showing that differences in firm pay premia contribute to earnings gaps along other dimensions, such as gender (Card et al., 2016) and race (Gerard et al., 2021). For example, Gerard et al. (2021) find that skill-based sorting contributes about 55-65% of the firm pay gap between whites and non-whites in Brazil. While the approaches differ, our estimates of the role of skill sorting for differences in firm pay by family background in Sweden are thus rather similar. Several recent studies highlight the the extent to which labor market sorting contribute to cross-sectional inequality (e.g. around 20% in Card et al., 2013 or 11-12% in our data). Our findings suggest that differences in firm pay contributes a considerably higher share to *intergenerational* persistence in earnings.

1 Data and specifications

1.1 Data

Our empirical analyses require data on earnings, employers, education, and other background characteristics of parents and their children covering multiple decades. For this purpose, we combine a set of linkable administrative registers made available by Statistics Sweden. We use tax registers for earnings records, full-count employer-employee data to identify the firms and establishments of workers, the education register for highest level and years of education, and the multigenerational register to link children to parents.

The earnings data cover the full working population for the years 1968-2018. Our main analyses use gross annual earnings from work including self-employment, bonuses and fringe benefits, and short-term (employer provided) sickness benefits.[1] Firm and establishments including their location and industry affiliation are available for the years 1985-2018. Official birth and family registers allow us to match nearly all Swedish-born children born 1932 or later, and foreign-born children born 1961 or later, to both their parents.

Main intergenerational sample. Our main analysis is based on men and women born 1967-1977, who can be observed on the Swedish labor market between age 25-41, and have at least one observed firm fixed effects at those ages. We focus on these cohorts such that we can observe a long and important part of their own labor-market career, while still being able to observe the prime-age incomes of their fathers. We measure the father's long-run income between ages 45 and 55, and consider as long-run income measures either the mean log earnings or mean earnings rank across the 45-55 age range. Fathers annual earnings observations are net of year dummies and quadratic age effects and ranks are computed relative to fathers in the main sample.We drop individuals for whom we cannot observe their father's income, which only applies to a small fraction of individuals (typically for migration-related reasons). In parts of our analysis we focus on the child generation's peak career outcomes, measuring earnings and employers at age 39-41. About 10% of the main sample drops out due to missing firm or earnings at these ages.

AKM sample. To estimate firm premia using an AKM model we use a second matched employer-employee sample, covering the entire Swedish labor force age 20-64 in the period 1985-2018. For this sample we use information on annual earnings, firm, age, gender, and education. The estimated firm and individual fixed effects are then used as inputs in our analyses of the main intergenerational sample.

Descriptive statistics. Table 1 provides descriptive statistics of the AKM sample and our main intergenerational samples. Differences between these samples are due to differences in age composition and due to the AKM samples containing observations from earlier years. As individuals in the AKM sample are on average born 12 years earlier, the fraction with a college degree and their log earnings are slightly lower compared to the main samples. The average log firm size is 6.35, but our data also allows us to identify the worker's specific establishment within large firms.[2] Our full-population data contain 1,301,551 observations

[1] In all analyses we exclude very low annual earnings observations. For each year, we compute the median earnings of men aged 45, and set annual earnings observations corresponding to less than 20% of this median to missing. This ensures that our estimates are not overly influenced by variation in labor supply.

[2] Average firm size is relatively large for a few reasons. First, Sweden has a large manufacturing sector with many big firms. Second, public sector workers are included in our samples, and all working for a certain municipality count as working for the same "firm". Third, small firms without sufficient firm-to-firm mobility

with positive earnings at age 39-41. Imposing non-missing father links (e.g. due to migration) and earnings of the father, this number drops to 1,076,969 and 1,059,546, respectively. The sample size further drops to 910,665 when excluding very low earnings observations (of the child or father). Lastly, requiring a valid firm connection with identified firm fixed effect (and demographic characteristics such as education) leaves us with an age 39-41 sample of 857,064 observations.

Table 1: Descriptive statistics

	AKM sample	Main sample (born 1967-77)	
Age	20-64	25-41	39-41
Earnings years	1985-2018	1992-2018	2006-2018
Log earnings	12.46	12.54	12.79
Share women	0.49	0.50	0.49
Share with college degree	0.38	0.47	0.49
Year of birth	1960	1972	1972
Log firm size	6.35	5.97	6.05
Log establishment size	4.21	4.16	4.18
Number individuals	7,668,377	967,417	857,064
Number firms	341,798	228,285	118,258

Notes: Descriptive statistics for different samples. Column 1 shows statistics for the AKM sample, covering individuals aged 20 to 64 born between 1922 and 1997. Columns 2 and 3 show statistics for the intergenerational sample born between 1967 and 1977, separately for the 25-41 and 39-41 age ranges.

1.2 Estimation of Worker and Firm Fixed Effects

We use the widely applied two-way fixed effects framework of Abowd et al. (1999) to decompose earnings into firm and individual components, conditional on a set of time-varying controls. Our specification for the log earnings, y_{ijt}, of individual i in year t while employed in firm $j = J(i,t)$ is:

$$y_{ijt} = \alpha_i + \psi_j + \boldsymbol{X}_{it}\boldsymbol{\delta} + \varepsilon_{ijt} \tag{1}$$

where α_i is a worker fixed effect, ψ_j a firm fixed effect ("firm pay premium"), $\boldsymbol{X}_{it}$ a vector of time-varying controls with coefficient vector $\boldsymbol{\delta}$, and ε_{ijt} an error term. The time-varying covariates in $\boldsymbol{X}_{it}$ include year dummies and a restricted set of education-gender-specific age dummies. Due to the well-known collinearity between age, cohort, and time, unrestricted age dummies are not identified. Rather than imposing a particular functional form for lifecycle earnings profiles, we follow Engbom et al. (2023) and assume that the age effects are constant for ages when the earnings profile is roughly flat, where we impose the age effects to be constant between ages 45 to 54. For studying lifecycle firm sorting by parental income

are excluded from all samples (see below).

(Section 3) it is important to account for education-specific variation in earnings over age, given that own education and parental income tend to be positively correlated. We therefore allow for earnings profiles to vary freely by education groups.

The firm pay premia ψ_j in equation (1) are identified from conditional changes in earnings as workers switch firms. In particular, they are only identified relative to some baseline firm within a set of firms connected through such firm-to-firm transitions ("movers"). Firm fixed effects identified within non-overlapping connected sets of firms are not directly comparable to each other. We therefore follow standard procedures to compute the largest connected set for our time period and drop worker-year observations for which workers are employed in firms outside of this set (approximately 0.7% of all worker-year observations). Firm fixed effects tend to be noisily estimated for firms connected to other firms by only a small number of movers (e.g. Andrews et al., 2008; Bonhomme et al., 2023; Kline et al., 2020), which leads to an inflated variance of the firm component in a variance decomposition based on equation (1). We thus drop firms (and associated worker-year observations) that are connected to other firms through fewer than five movers (an additional 6% of worker-year observations). [3] This adjustment together with the fact that we use population-wide data for a long time period should diminish concerns about such limited-mobility bias.

We estimate equation (1) using our employer-employee sample for the years 1985-2018. We focus on full-time workers, approximated by excluding worker-year observations with annual earnings lower than 20% of the yearly median earnings of 45-year old males. We also provide robustness tests using data on actual wage rates instead of annual earnings for a large subsample covering the same time period (see Appendix A1.2). In the subsequent sections, we use our estimates of α_i and ψ_j as inputs in our intergenerational analysis.

Using our estimates from equation (1), we can decompose the variance in earnings into its different components; Appendix A1.1 reports such decompositions for both the AKM and the intergenerational samples. We find that worker effects explain 29-38% while firm effects explain 7-11% of the variance in log earnings. The covariance between worker and firm effects contributes another 7-11%, reflecting the sorting of more productive workers to better-paying firms (assortative matching). Overall, our decomposition results are similar to Engbom et al. (2023), who use similar data and specifications, and also largely in line with evidence from the US (e.g. Song et al., 2019).

2 The contribution of firms to intergenerational persistence

We start by providing some basic facts on how children from different socioeconomic backgrounds sort into firms with different pay premia, as well as how such firm sorting contributes

[3] However, table A4 in appendix A1.2 shows that the results remain similar when we do not exclude firms with fewer than 5 movers.

to intergenerational income persistence. To this end, we estimate variations of the linear regression

$$y_{ijt} = \alpha + \beta y_{f(i)} + u_{ijt}, \tag{2}$$

where y_{ijt} is child log earnings, $y_{f(i)}$ the log earnings of the father of child i, and β is the intergenerational earnings elasticity (IGE) .We measure the father's long-run income as the mean residual of log income for the ages 45 and 55, where log income has been residualized of year dummies and quadratic age effects. The childrens income are measured as mean of log income for the ages 39-41. As y_{ijt} can be decomposed according to equation (1), the slope coefficients from separate regressions of each of its components on $y_{f(i)}$ will together sum up to β. Our primary focus is on the slope coefficient from a regression of the child's estimated firm pay premium $\hat{\psi}_j$ on $y_{f(i)}$, which we denote β_{firm} and refer to as the *socioeconomic status gradient in firm pay* or, shorter, *SES gradient in firm pay*.

Table 2 shows estimates of the IGE and its components. The estimates are based on our main intergenerational sample, with earnings and firm premia measured as averages over age 39-41. As shown in column (1), the IGE is roughly 0.20 in our sample, which aligns with prior Swedish estimates of the IGE in *labor* income in pooled samples that include both sons and daughters (e.g. Brandén and Nybom, 2019; Engzell and Mood, 2023).[4] Columns (2) and (3) of Table 2 decompose the IGE into individual and firm components. About 60% of the IGE is attributed to the individual effects, which capture all permanent determinants of earnings (e.g. time-constant abilities and skills). The firm effects contribute 27% of the IGE, rising to 31% if we first net out the AKM covariates and residuals from y_{ijt}.[5] The firm pay gradient explains therefore a sizable share of income persistence from one generation to the next.[6] Its magnitude for Sweden contrasts with earlier work by Dobbin and Zohar (2023), who find that firms contribute about 22% to the IGE (after having netted out the AKM residuals) in Israel – about two thirds of the share that we find in our setting. The firm contribution we find is also slightly larger than in Engzell and Wilmers (2024), who use similar Swedish data but focus on percentile ranks of earnings and firm pay premia.

Why do we find stronger sorting of workers across firms by family background in Sweden than previous work from Israel? One possibility is that firm sorting is really stronger in Sweden. But another possible explanation is that differences in the quality of the underlying income data explain part of this gap. In particular, while Dobbin and Zohar observe all workers in the Israeli labor market in a six-year period (2010-2015), we observe the entire Swedish labor force over a 34-year period (1985-2018), allowing us to pin down firm pay

[4]The IGE tends to be slightly higher for total income, especially for sons (e.g. Nybom and Stuhler, 2016). In our sample, the IGEs estimated separately for sons and daughters are 0.23 and 0.17, respectively.

[5]The correlation between the log of father's income and $\boldsymbol{X}_{it}\boldsymbol{\beta}$ and ε_{ijt} explains 14% of the IGE (see column 4). Abstracting from this component, firm effects explain 27%/(59%+27%)=31% of the IGE.

[6]As shown in Appendix Table A6, the firm pay gradient is slightly stronger for women than men.

Table 2: Decomposition of the intergenerational earnings elasticity

	Dependent variable				
	y_{ijt} (1)	$\hat{\alpha}_i$ (2)	$\hat{\psi}_{j=J(i,t)}$ (3)	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$ (4)	$\hat{\psi}_{j=J(i,t)}$ (5)
$y_{f(i)}$	0.201***	0.118***	0.054***	0.029***	0.037***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
$\hat{\alpha}_i$					0.151***
					(0.001)
Share of IGE	100%	59%	27%	14%	18%
Worker obs.	857,064	857,064	857,064	857,064	857,064

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings at ages 39-41 on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child log earnings y_{ijt} according to equation (1) into individual fixed effects α_i, mean firm fixed effects ψ_j for the same ages earnings are estimated, and time-varying controls. Robust standard errors in parentheses.

premia more precisely..When we replicate their data scenario, with an earnings panel over the years 2010-2015, we find a substantially diminished SES gradient in firm pay (see Appendix Table A2), explaining only 16% of the IGE (18% when netting out the AKM covariates and residuals). A classical measurement noise should not effect the estimate since firm pay is the dependent variable. However, figure A1 shows that the noise it not classical, instead the firm premia estimated for the short period seems to systematically overestimate the firm premia at low values of the firm premia and underestimate the value of the firm premia for high values of the firm premia. Differences in the quality of the AKM estimates therefore seems to affect the estimated firm contribution to the IGE.[7][8] Section 2.3 provides more evidence on the sensitivity of our estimates to sampling and specification choices.

Skill Sorting. While worker-firm sorting thus explains a substantial share of the IGE, such sorting might be only indirectly related to family background. Prior research has documented assortative matching between workers and firms, such that firm and individual fixed effects tend to correlate positively. This sorting of more productive workers to higher-paying firms contributes to income inequality in the cross-section, as shown in eq. (A1) and Appendix Table A1 (see Card et al., 2013; Song et al., 2019). In addition, worker-firm sorting increases intergenerational persistence, since individual fixed effects correlate positively with parental

[7] While the results from the measurement error correlations are in line with the importance of high quality data to estimate the AKM, differences in firm pay for the specific periods could also potentially explain the results.

[8] The quality of the data differs also in other aspects. First, we measure parental earnings over a longer age span, between age 45-55. Second, we can retain 93% of all Swedish children born in 1967-1977 in our analysis, while Dobbin and Zohar (2023) drop about half of their sampled cohorts due to insufficient information on earnings or incomplete sets in the AKM estimation. We show in Appendix Table A3 that randomly dropping half of our sample has little effect on our estimates. Excess homogeneity in non-representative sample could bias estimates of the IGE downward (Solon, 1992), but it is less clear how it would impact the share of the IGE attributable to firms.

income (see column (2) of Table 2). The SES gradient in firm pay could therefore be a "mechanical" consequence of the assortative matching between workers and firms as documented by prior research.

To abstract from the general degree of assortative matching between workers and firms, we follow Dobbin and Zohar (2023) and regress the firm fixed effect on the log of father's income *conditional* on the individual fixed effect (column (5) of Table 2). The coefficient on the individual fixed effect is large and statistically significant, reflecting assortative matching. The coefficient estimate for log of father's income now falls by about 30%, to 0.037. Part of the SES gradient in firm pay thus originates from gradients in (permanent) worker-level characteristics (e.g. skills), as captured by the individual fixed effects. Put differently, by amplifying the pay-off to skills, worker-firm sorting magnifies the well-known SES gradient in skills and human capital. According to the estimates here, this mechanism explains about a third of the firm contribution to intergenerational persistence; even after accounting for assortative matching, systematic differences in firm premia still explain roughly 18% of the IGE. The sorting of workers across firms by family background appears therefore substantially stronger than one would expect based on the observed degree of skill sorting in a worker-firm fixed effects model. However, skill sorting is difficult to measure, and we return to this question in Section 4.

2.1 Non-linear firm pay gradients

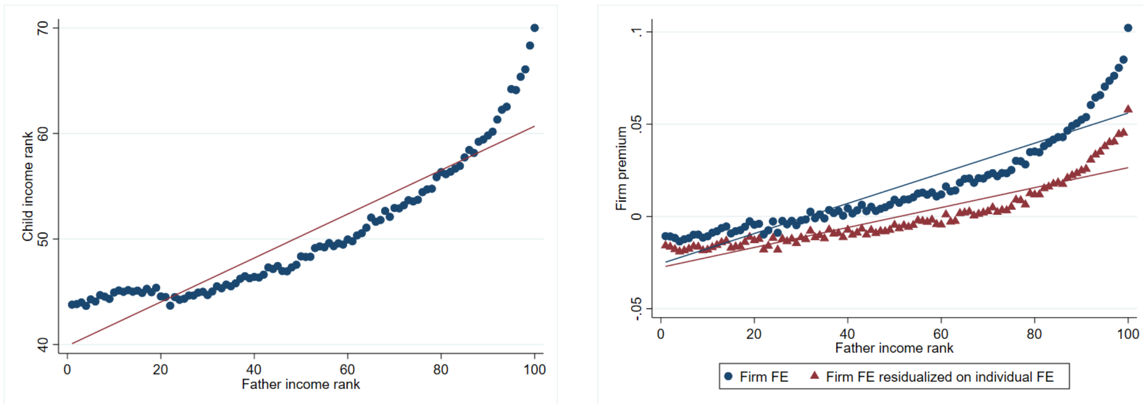
Does the sorting across firms matter more among low- or high-income families? Figure 1 shows how the expected child earnings (subfigure A) and firm premia (subfigure B) at age 40 vary across the distribution of parental income, with incomes now expressed in ranks (see the corresponding Figure A2 in the Appendix for log income). The two figures reproduce the positive relationships from Table 2, but with a gradient that is strongly increasing starting from about the 75th percentile of the parental-income distribution. The SES gradient in firm pay is qualitatively similar when conditioning on the individual fixed effect (red triangles in subfigure B), but smaller in magnitude. Moreover, the difference between the unconditional (blue circles) and conditional (red triangles) gradients increases along the distribution, implying that skill-based sorting is particularly important among richer families. Overall, the results show that sorting into high-paying firms is an especially important driver of intergenerational transmission among high-income families.

2.2 The roles of region, industry, and firm size

We next explore some potential determinants of the SES gradient in firm pay. Pay premia differ not only across firms, but also across regions and industries. For example, the industries

Figure 1: Child income and firm premia by father's income rank

(a) Child income rank (b) Firm premia at age 40



Notes: Sub-figure (a) shows binned scatter plots of child's income rank at age 40 by father's income rank. Sub-figure (b) shows firm fixed premia ψ_j at age 40 as estimated by equation 1 and firm premia residualized on individual fixed effects by father's income rank.

with the highest mean firm pay premia are oil and natural gas extraction, financial sector support services, telecommunications manufacturing, chemical manufacturing, IT services, and research and development. Moreover, large firms tend to pay higher premia than smaller firms. The positive relationship in Table 2 could therefore, among other things, be due to children entering firms in similar industries as their fathers, or located on the same local labor markets.

Table 3 extends our prior analysis by controlling for various sets of fixed effects. Controlling for region (21 counties, column 2) diminishes the SES gradient in firm pay by roughly 20%. Redoing the main analysis within 2-digit industries (59 in total, column 3) instead diminishes the SES gradient by about 36%. Controlling for the size of the workplace has a smaller, yet non-trivial effect on the estimated SES gradient (column 4). Controlling for region, industry and size jointly has a noticeable impact, with only about 42% of the unconditional SES gradient in firm pay being left unexplained. Thus, an important share of the SES gradient can be traced to observable firm characteristics, though a similarly important part of the gradient remains unaccounted for.

2.3 Sensitivity to specification and sampling choices

We perform a number of tests to probe the sensitivity of our estimates to specification. First, one might worry that the time period we use to estimate the AKM components is *too* long; while we capture a lot of worker flows between firms, the assumption that firm-specific pay is fixed over time is less likely to hold over long time spans. Thus, we consider a modified AKM specification in which firm pay is allowed to vary over time (e.g. Lachowska et al., 2020; Engbom et al., 2023). To estimate the time-varying AKM we divide the period 1985-

Table 3: Decomposing the relationship between firm premia and parental income

	Dependent variable: Estimated firm pay premium $\hat{\psi}_{j=J(i,t)}$				
	(1)	(2)	(3)	(4)	(5)
$y_{f(i)}$	0.053***	0.041***	0.034***	0.045***	0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Region FEs		X			X
Industry FEs			X		X
Establishment size (log)				X	X
Share of β_{firm}	100%	77%	64%	85%	42%
Worker obs.	812,697	812,697	809,758	780,025	779,596

Notes: Column (1) reports estimates of the slope coefficient from regressing $\hat{\psi}_j$, as estimated from equation (1), on log fathers earnings. Columns (2)-(5) report coefficient estimates from the same regression controlling for region fixed effects (21 counties), industry fixed effects (2-digit level, 59 industries), or log establishment size. Robust standard errors in parentheses.

2018 into 4 periods, and allow the firm fixed effects to vary between the periods. As shown in Appendix Table A4, Panel A, this modified specification only renders a very marginal increase in the role of firms; we therefore continue to use the traditional AKM specification as our baseline. Second, we consider sorting across *establishments* rather than firms. As shown in Panel B of Table A4, establishment fixed effects explain 35% of the IGE (39% if we first net out the AKM residuals), a substantially greater share than firm fixed effects (cf. Table 2). While there are arguments in favor of using establishment codes – for example, public sector employers constitute large and somewhat artificial "firms" – we align with the bulk of the related literature and stick to using firms in our main analyses. Third, we explore whether the results are sensitive to using wages rather than annual labor earnings as outcome. Wages are observed for a subsample of workers in an employer survey. While the role of firms is considerably diminished (see Appendix Table A5), part of this decrease appears to be due to the wage sample being smaller and more selective (the wage survey oversamples workers in large firms and the public sector). We thus proceed with using earnings, as much of the prior literature, but recognize that variation in labor supply might contribute to the estimated AKM components.

We further test how our estimates vary with sampling choices. One question is whether the firm pay gradient is driven by children choosing to work in the same firm as their parents. However, while the share of "firm followers" is indeed non-negligible in our sample (in line with evidence by Corak and Piraino 2010, Stinson and Wignall 2018), the estimated firm pay gradient remains nearly unchanged when we drop them from the estimation sample (Appendix Table A6, Panel C). In contrast, dropping public sector workers raises the firm contribution by 10% (Table A6, Panel D). This might reflect that wages are more compressed in the public sector, but also that some public employers are very large, such that their firm fixed effects may not well reflect pay in different groups within the "firm". Finally, restricting the sample to firms for which we observe at least 10 or 50 "movers" in the AKM estimation decreases the firm pay gradient (Table A6, Panels E and F). On the one hand, dropping small firms with few movers decreases the limited mobility bias in AKM estimates (Bonhomme et al., 2023). On the other hand, firm fixed effects might be less informative about worker-specific pay in large than small firms. We therefore stick with our baseline specification, which retains firms with at least five movers.

3 Firm pay gradients over the lifecycle

The previous section showed that children from high-income families end up working at higher-paying firms at age 40, that this pattern holds conditional on their own permanent skills as captured by the individual fixed effects, and that this firm-sorting explains a sizable part of the persistence of income inequality across generations. This section explores the career dynamics of firm pay, with the aim to further understand how and why children from more privileged family backgrounds end up at higher-paying firms.

3.1 Firm pay premia over the lifecycle

Figure 2 plots age profiles of the mean firm pay premium by quartile of parental income. Subfigure (a) shows a strong SES gradient in firm pay already at age 25, with children from higher parental-income quartiles working at firms with higher pay premia. Further, children from the top parent quartile (yellow squares) improve their firm premia at a substantially faster rate than other children in the early part of their careers: while the bottom three quartiles increase their firm premium by about 1 percentage point up until the early 30s, the increase among children from the top quartile is about 3 percentage points. The gap in firm pay stabilizes in the mid-30s. The standard deviation of the firm premium is 0.14-0.15, depending on the sample (see Table A1), which implies that children in the top quartile have a firm pay advantage compared to the next quartile corresponding to about 29% of a standard deviation.

However, some of these dynamics might be driven by education-specific differences in career profiles and systematically higher education levels and delayed labor-market entry among children from well-off backgrounds. Subfigures 2b and 2c thus reproduce the same plot separately for children with at most high school and some college or more, respectively. In both groups, a similar type of SES gradient in firm pay is apparent. However, among non-college children, firm premia generally grow less over age. Among college children, all quartiles grow their firm premia over age, but the top-quartile children do so at a faster rate than others. While the levels and early-career growth in firm premia are higher among college graduates, the qualitative patterns are surprisingly similar whether we consider all children or only children with similar levels of education.[9][10]

Given the stability of the patterns when comparing the full sample and education-based subgroups, we proceed to focus on the former. However, results by education are reported in Appendix A3, and we discuss noteworthy education-related differences in the main text. Finally, Figure A4 in the Appendix plots the difference in firm pay premia over the lifecycle conditional on individual fixed effects, to account for the general degree of assortative matching between workers and firms that is due to skill sorting (i.e., the lifecycle version of column (5) of Table 2). The gaps in firm pay premia as well as the large early-career increase in firm pay among children with high-income fathers remain largely similar.

In sum, gaps in firm pay open up already at the beginning of the career, with high-SES children being substantially more likely to work in firms with more generous pay policies. This gap increases further in the early career stages, and then stabilizes by the mid-30s. In Section 4 we show that the initial gap *cannot* be explained by differences in worker skills, while the subsequent widening of the gap is mainly explained by skill differences that correlate with parental background.

3.2 Climbing the job ladder

We showed that high-SES children not only start their career at higher-paying firms, but also that a large part of the long-term SES gradients in firm pay build up over the first 10-15 years on the labor market. This divergence can be due to either *more frequent* firm switches among high-income children (i.e., climbing the job ladder faster), or that they do *better* switches (i.e., the rungs of their ladder are further apart), or a combination of both.

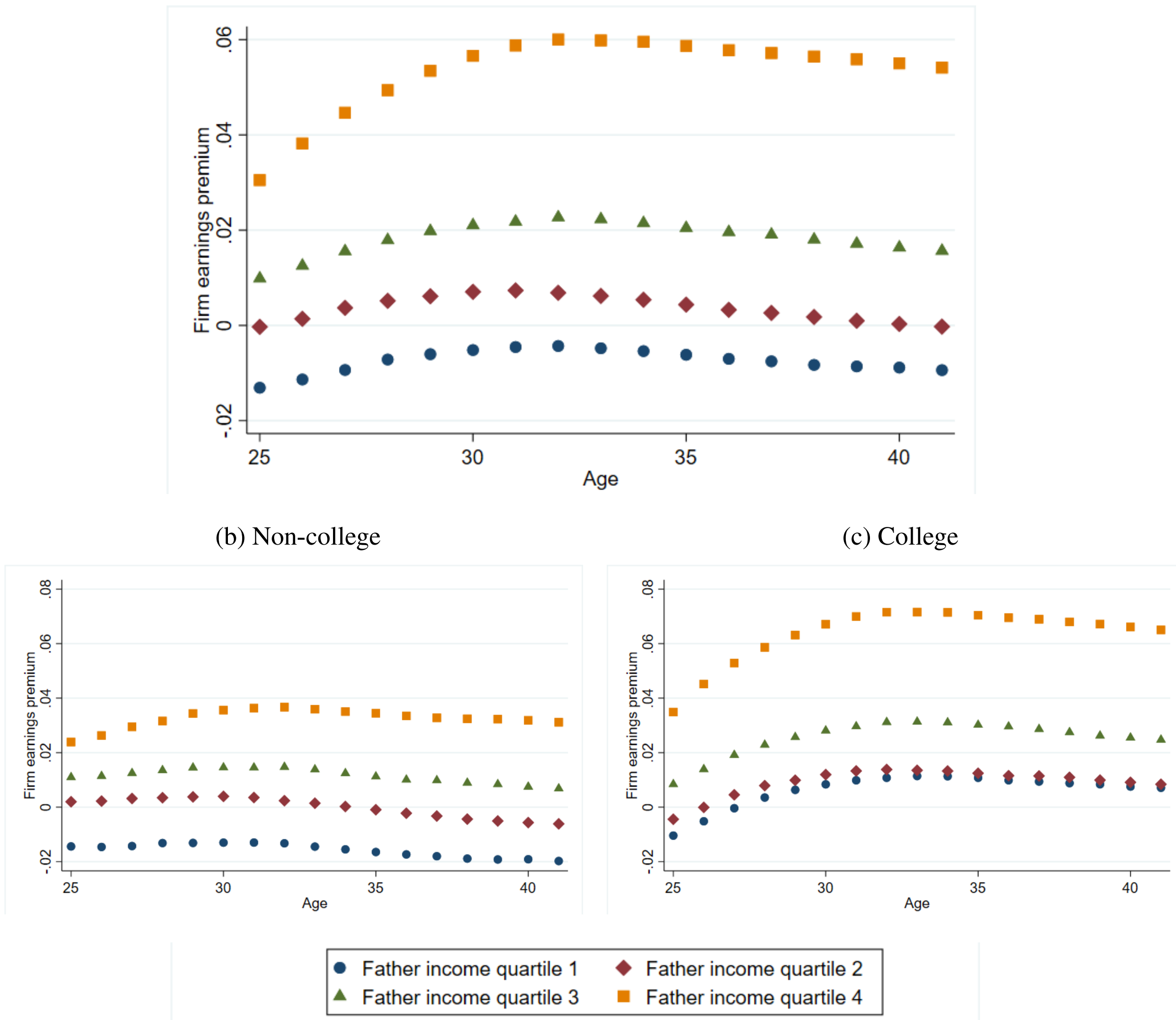
Figure 3a shows the annual firm switch rate, i.e. the probability of being observed in a new firm at age a compared to at age $a - 1$, by age and parent-income quartile. As expected, all children are more likely to switch firms at early age, with roughly 20-25% of young workers annually switching firm through ages 25-30, while only 11-12% switch at age 41. Up until the early 30s, there is also a small but clear SES gradient in the likelihood to switch firm, with high-SES children being more avid switchers. These differences in the probability of switching diminish over age, and become negligible after age 35. The patterns hold within education groups (Appendix Figure A6), though overall higher education is associated with slightly more firm switching and the elevated switching of high-SES children is more pronounced in the college group.

[9] This similarity reflects that time-constant differences in worker pay are captured by the individual fixed effects, while differences in lifecycle growth within firm are captured by the education-age interactions in equation (1).

[10] To further probe that our results on career dynamics are not driven by differences in age at labor-market entry we reproduce the analysis by potential , finding very similar results (see Appendix A3).

Figure 2: Firm earnings premium over the lifecycle

(a) Full sample



(b) Non-college

(c) College

Notes: The figures plots the mean estimated firm premium $\hat{\psi}_j$ over the life cycle, by quartile of father's income. Sub-figure (a) shows the result for the full sample, sub-figure (b) shows the result for children without collage education and sub-figure (c) shows the result for children with college education. Father's income quartiles are defined in the full sample.

Simply moving to a new firm is not enough to improve your firm pay, unless such moves entail improvements in firm pay premia.[11] Figure 3b plots the proportion of firm switches that are indeed premium-improving, defined as moving to a firm at age a with an estimated firm premium that is strictly larger than the one in the previous firm at age $a-1$, again computed separately by age and parent-income quartile. Children tend to switch to better firms at early age (i.e., they "climb the firm ladder"), as the chance of a "good" (premium-improving) firm switch is greater than 50% for all SES groups. But again there are distinct differences with respect to SES, with high-SES children being more likely to experience such premium-improving switches. The likelihood of good and bad switches tend to even out over age,

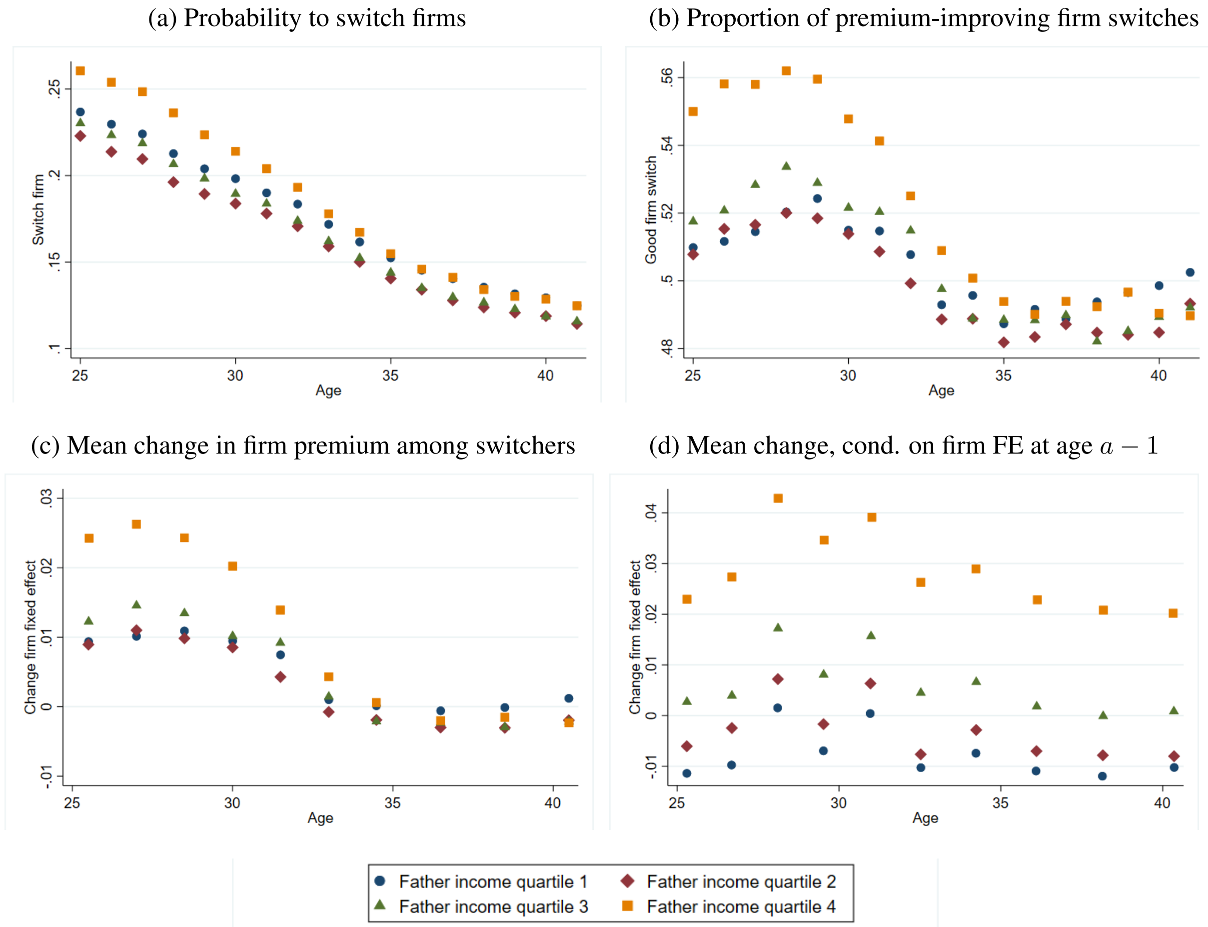
[11] While there are other potential reasons why worker mobility may benefit human capital and wage growth, we focus here on the benefits in terms of firm pay premia.

as well as the SES gradients. Moreover, while low-SES children are more likely to switch firms than mid-SES (subfigure a), they are not more likely to gain in firm pay conditional on switching. This observation indicates that high- and low-SES children switch for different reasons, e.g. the latter might to a larger extent be involuntary switches or moves between various forms of temporary employment.

Figure 3c shows the average change in the firm premium in a similar fashion. The figure shows a clear gradient indicating that with every firm switch, high-SES children raise their earnings by about one additional percent compared to others through improved firm pay. Put differently, children from the top quartile of parental income enjoy gains in firm pay that are more than twice as large as the gains achieved by children in the lower quartiles. This difference shrinks over age, and after age 35 the effect of switches on firm pay premia is close to zero, independent of parental income. However, this closing of the gaps at later ages may be a mechanical consequence of gaps at earlier age: rapid improvements at early age should mechanically make it harder to improve your firm premium at later age, while those who remain working at low-pay firms at early age retain more scope for improvements in firm pay later on. In the extreme scenario of fully randomized switches, low-SES children with lower initial firm premia would experience much larger gains from switching.

To account for this mechanical relation between the change and level of firm pay premia, Figure 3d reports the change in firm premia while conditioning on the level of the firm fixed effect in the previous year. Thus, instead of relating the new firm premium to your own premium in the year prior, we implicitly relate them to the mean premium of the cohort in the year prior. We find that over the entire age range, high-SES children switch to better firms than children in the lower quartiles. Thus, the estimates in Figure 3c are partly compressed by the fact that premium-improving switches are increasingly hard to find for those who have already climbed high up on the firm ladder. Accounting for this mechanical relation, the SES gradient is now essentially constant over the lifecycle.

Figure 3: Firm switching patterns over the lifecyle



Notes: The figures show pattern for switching firms over the life-cyle, by quartile of father's income. Sub-figure (a) shows the probability of working in a new firm compared to the year before. Sub-figure (b) shows the probability of switching to a firm with a higher firm premium than the one before conditional on switching. Sub-figure (c) shows the difference between the new firm premia and the firm premia before for individuals who switch firms. Sub-figure (d) shows the same but adds a control for the firm premia in the firm before the switch.

Appendix Figures A7 and A8 show the proportion of "good" firm switches and the mean change in firm pay conditional on switching separately by education group. College-educated individuals are generally more likely to experience premium-improving switches and larger boosts in firm pay throughout the career, but the SES gradients are otherwise similar across education groups. Moreover, we show that the larger gains in firm pay among high-SES children does not appear to be driven by differences in unemployment – even considering only voluntary switches without any intermediate unemployment there is an SES gradient in the quality of switches.[12]

[12] We can use data on UI benefit receipts to distinguish voluntary switches (without any UI receipt in between employment spells) and involuntary switches via spells of unemployment (as measured by UI benefit receipt). As shown in Figure A9, high-SES children are more likely to experience improvement in the firm premium following both voluntary and involuntary switches.

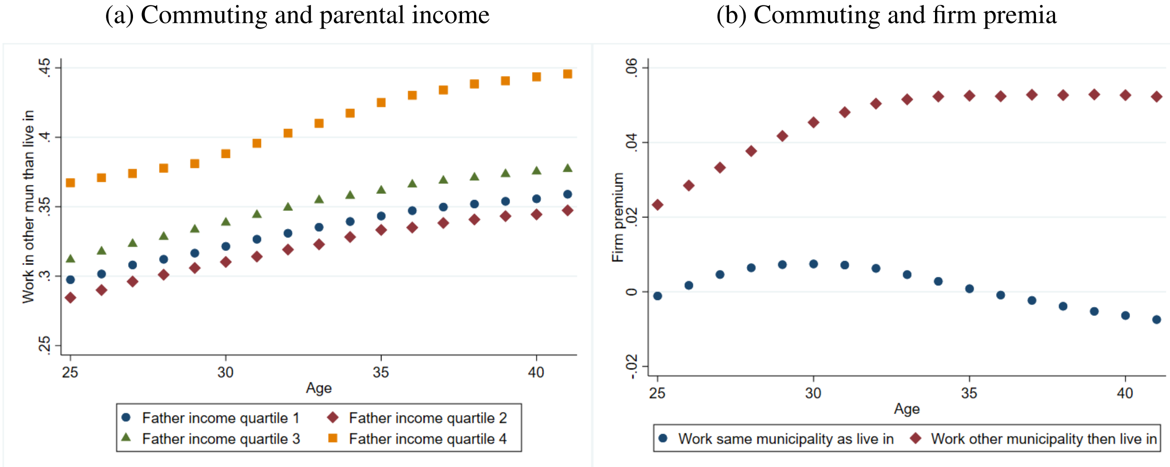
3.3 Commuting up the job ladder?

One reason for why children from high-income families gain higher firm pay premia over the life-cycle is that they might be more likely to commute longer for work. In particular, they might have access to wider social networks and/or have access to support structures that make commuting easier, such as child care or different modes of transportation. By expanding the choice set of feasible jobs, commuting may increase wages and firm pay premia (Le Barbanchon et al., 2020; Agrawal et al., 2024). On the other hand, commuting is costly, both financially and in terms of time. Hence, if differences in commuting patterns can explain a sizable fraction of the SES gradient in firm pay, then gaps in firm pay premia might overestimate the implied gaps in overall welfare.

Figure 4a shows that children from high-income families are indeed more likely to commute, which we define as working and residing in different municipalities. At age 25, the share of commuters is 8 pp. (nearly 30%) higher among children in the top compared to bottom quartile of parental income at age 25, increasing to 10 pp. by age 40. Figure A11 in the Appendix shows that these findings also hold within education groups; while more educated workers are more likely to commute, differences in education do not explain the strong SES gradient in the commuter share.

Figure 4b shows that commuters earn higher firm premia, in particular at later ages. At age 40, the gap in firm premia between commuters and non-commuters reaches 6 pp., nearly half of a standard deviation. However, this gap also reflects selection effects, as workers who commute might differ in other important ways from those who do not (e.g., live in a bigger city, where it is more common to work and reside in different municipalities). To abstract from selection, Table A9 in the Appendix reports event-study type regression estimates conditional on individual fixed effects, showing how firm pay changes for individuals who begin to commute. We find that commuting raises average firm premia "only" by about 1.7 pp. at age 40. The SES gradient in the probability to commute, although considerable, contributes therefore little to the SES gradient in firm pay.

Figure 4: Commuting



Notes: Sub-figure (a) shows the proportion of individuals who commute (i.e., work in another municipality than they live in) over the life-cycle, by quartile of father's income. Sub-figure (b) plots the mean firm premia by age and commuting status.

3.4 Beyond firm pay premia

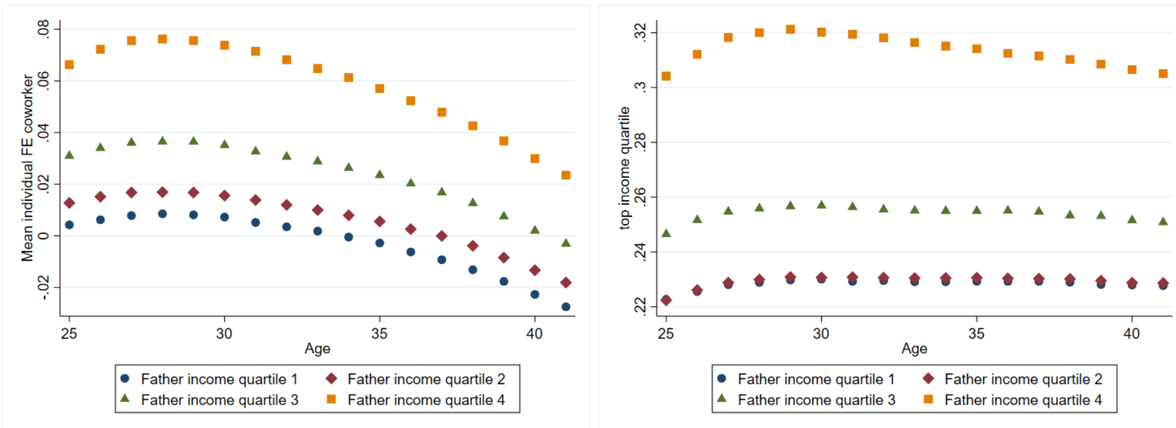
Firms differ in other important dimensions apart from their pay premia. In this section, we study whether the type of firms in which children from high-income families work offer other advantages apart from higher pay. We begin by considering "static" firm characteristics, such as the composition of the workforce. Figure 5a shows that children from high-income families tend to have more productive co-workers, as captured by their estimated AKM worker fixed effects $\hat{\alpha}_i$. The mean co-worker effect is about 6 pp. higher for children from the top compared to bottom quartile of parental income. They are also exposed to a higher share of co-workers from high-income families, as shown in Figure 5b. This type of firm-level segregation is relatively stable over age – if anything, workers are more segregated by own productivity and SES at early than at later age. If working with more productive co-workers increases one's own productivity (as in Kremer, 1993), such segregation might also contribute to the firm pay advantage that we documented in Figure 3.1.

The fact that high-SES children end up in firms with more productive workforces (as measured by co-worker individual fixed effects) might have interesting implications regarding their career trajectories, as there might be positive spillovers on their own productivity due to learning or fostering of valuable social networks. Moreover, firms might offer different opportunities for career development, partly because different firms might themselves grow at different rates. As a consequence, firms may have different *dynamic* implications, apart from the static difference in firm pay as captured by the AKM approach.

To illustrate that firms also differ in a dynamic sense, Figure 6 shows the mean employment and earnings growth of co-workers over the following five years (i.e., between age a and age $a+5$). Figure 6a shows that children from high-income families are more likely

Figure 5: Static firm characteristics

(a) Individual fixed effects of co-workers (b) Share of co-workers with father in top quartile



Notes: The figures shows firm characteristics by age, by quartile of father's income. Sub-figure (a) shows the mean individual fixed effect of the co-workers. Sub-figure (b) shows the share of co-workers with fathers in the top income quartile.

to work in growing firms. Indeed, children from the top quartile of father's income work in firms that grow 50-100% faster than children in the lower quartiles.

High-SES children are also more likely to work in firms with high earnings growth, as shown in the next two sub-figures. Figure 7b plots the mean earnings growth between age a and age $a+5$ for co-workers who stay in the firm, while Figure 7c tracks the five-year earnings growth for co-workers irrespectively of whether they stay or leave the firm. Interestingly, the SES gradients are much more pronounced in Figure 7c, suggesting that much of the difference in co-worker earnings growth comes from mobility to new firms, rather than just differences in earnings growth for incumbent workers within firms. This finding is in line with the results shown above that children from high-SES families tend to work together, are more likely to switch firms, and make switches that render higher firm pay premia.[13]

3.5 Heterogeneity in firm-specific returns to experience

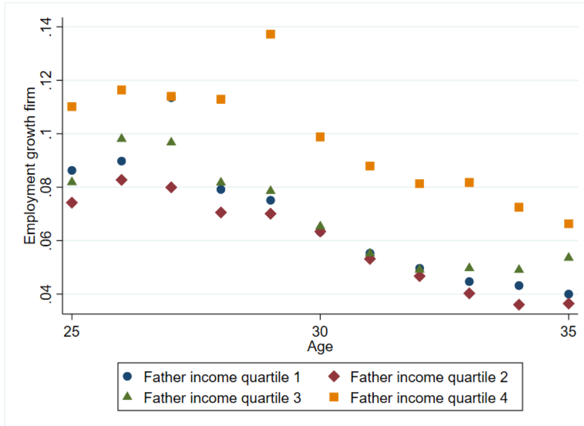
We found that children from high-income families are more likely to sort into growing firms and tend to have more productive co-workers, who enjoy faster earnings growth. These findings suggest that these firms may not only offer pay advantages in a "static" sense, but also better opportunities for career development in a "dynamic" sense. For example, faster earnings growth among co-workers could reflect better learning or promotion opportunities; and while some firms might not pay that well, it could still be optimal for individuals to work in such firms if they offer high returns to experience, increasing income later in the career.

To more explicitly estimate the role of such firm-specific returns to experience, we follow

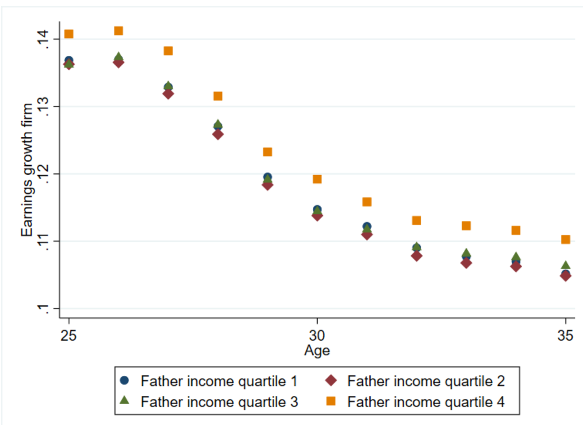
[13]Interestingly, the gaps between the top quartile of father's income and the rest in co-worker earnings growth and firm employment are relatively stable across age, while differences between those from the first, second, and third quartiles are generally much smaller.

Figure 6: Firm-level employment and earnings growth

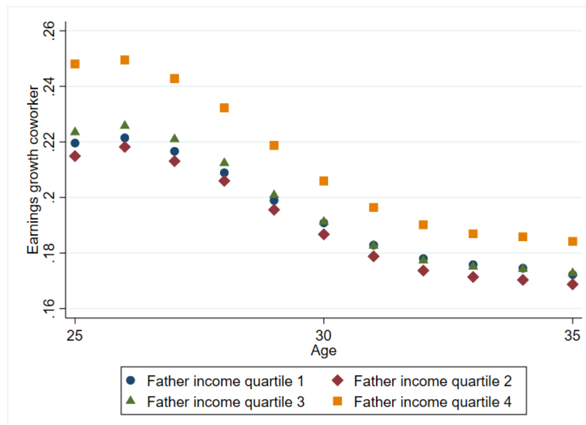
(a) Five-year employment growth of firm



(b) Five-year earnings growth of co-workers, stayers



(c) Five-year earnings growth of co-workers, all



Notes: The figures show the difference in firm characteristics between year t and year $t+5$ for the individuals who work in the firm at each indicated age. Sub-figure (a) plots the mean employment growth in the firm between year t and $t+5$, by quartile of father's income. Sub-figure (b) plots the earnings growth for coworkers who stay in the firm between year t and $t+5$. Sub-figure (c) plot the earnings growth for all coworkers, including both those who stay and those who switch firm between year t and $t+5$.

Arellano-Bover and Saltiel (2021) and split our sample of workers into two random samples. Using one of the random samples we divide firms into ten classes using the distribution of stayers' yearly unexplained earnings growth using a k-means clustering algorithm. We then use the other random sample to estimate the firm-class specific returns to experience following the methodology in Arellano-Bover and Saltiel (2021). In particular, we estimate an extended two-way fixed effects framework according to:

$$y_{ijt} = \alpha_i + \psi_j + \sum_{m=1}^{K} \gamma_m Exp(m)_{it} + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijt}, \tag{3}$$

where $Exp(m)_{it}$ is years of experience in firm class m up until year (or age) t. As above, we include individual and firm fixed effects, such that γ_m is identified from workers who are employed in the same firm, but have earlier experience from different firm classes.$\boldsymbol{X}\boldsymbol{\beta}$

controls for age and year fixed effects.[14]

Figure 7 summarizes our results. In sub-figure (a) we show the estimated returns per year of experience by firm class m, where class 1 is the firm class consisting of firms with the highest returns to experience and class 10 consists of the firms with the lowest returns to experience. Largely similar to Arellano-Bover and Saltiel (2021), we find important differences in the returns to working in firms belonging to different classes as defined by their unexplained earnings growth. However, the differences in firm-specific returns between the top and bottom classes are considerably smaller in our Swedish data compared to their data from Italy and (especially) Brazil: while workers in the top firms (firm class 1) experience an annual firm-specific boost to their earnings growth of 2.5 percentage points, the expected firm-related earnings growth in the bottom firms (type 10) is slightly negative. Note that these growth components are net of the general education-by-gender specific earnings growth component as captured by $\boldsymbol{X}_{it}\boldsymbol{\beta}$.

In Figures 7b and 8c we show how the probability to work in high-return firms (classes 1 or 2) and low-return firms (classes 9 or 10) differs by SES and over age. Children from high-income families are more likely to work in firms with very high returns to experience (belonging to the top-two classes) and less likely to work in firms with the lowest returns to experience (the bottom-two classes). The gap in the probability to work in firms with high returns is evident at all ages, and peaks around age 27-28, when about 16% of top-quartile children and about 10-11% of the non-top-quartile children work in high-return firms. Differences between those from quartiles 1-3 are generally smaller; if anything, it is those from the second quartile that have the lowest (highest) chance to work in firms with high (low) returns to experience.

Figure 8d shows the estimated average annual firm-specific return at the individual's current firm at different ages, separately by father's income quartile. We can see that top-quartile children are more likely to be employed in firms with higher returns to experience throughout the observed age range. There is also a clear tendency of an increasing SES gap in returns up until about age 28-29, while the gap shrinks after that age. In terms of magnitudes, those with parents in the top quartile enjoy firm-specific returns that are on average 20% higher between age 20 and age 30 than those with parents in the bottom quartile.

Table 3 quantifies how much of the intergenerational earnings elasticity at age 40 (column 1) can be attributed to the different components in equation (3). In particular, we decompose the contribution to the firm pay premium at age 40 into the early-career firm premium at age

[14] In contrast to the static AKM model, we here limit the sample to workers who we can track up to age 41, corresponding to cohorts born between 1967 and 1977 (i.e., the same cohorts as considered in our intergenerational regressions). Since we only include observations up to age 41 we cannot use the assumption that the effect of age on earnings is constant between ages 45-54, as we do in our main AKM specification to be able to include both age and year fixed effects. Instead we normalize age relative to age 40 and include second and third order polynomials of age interacted with education and gender.

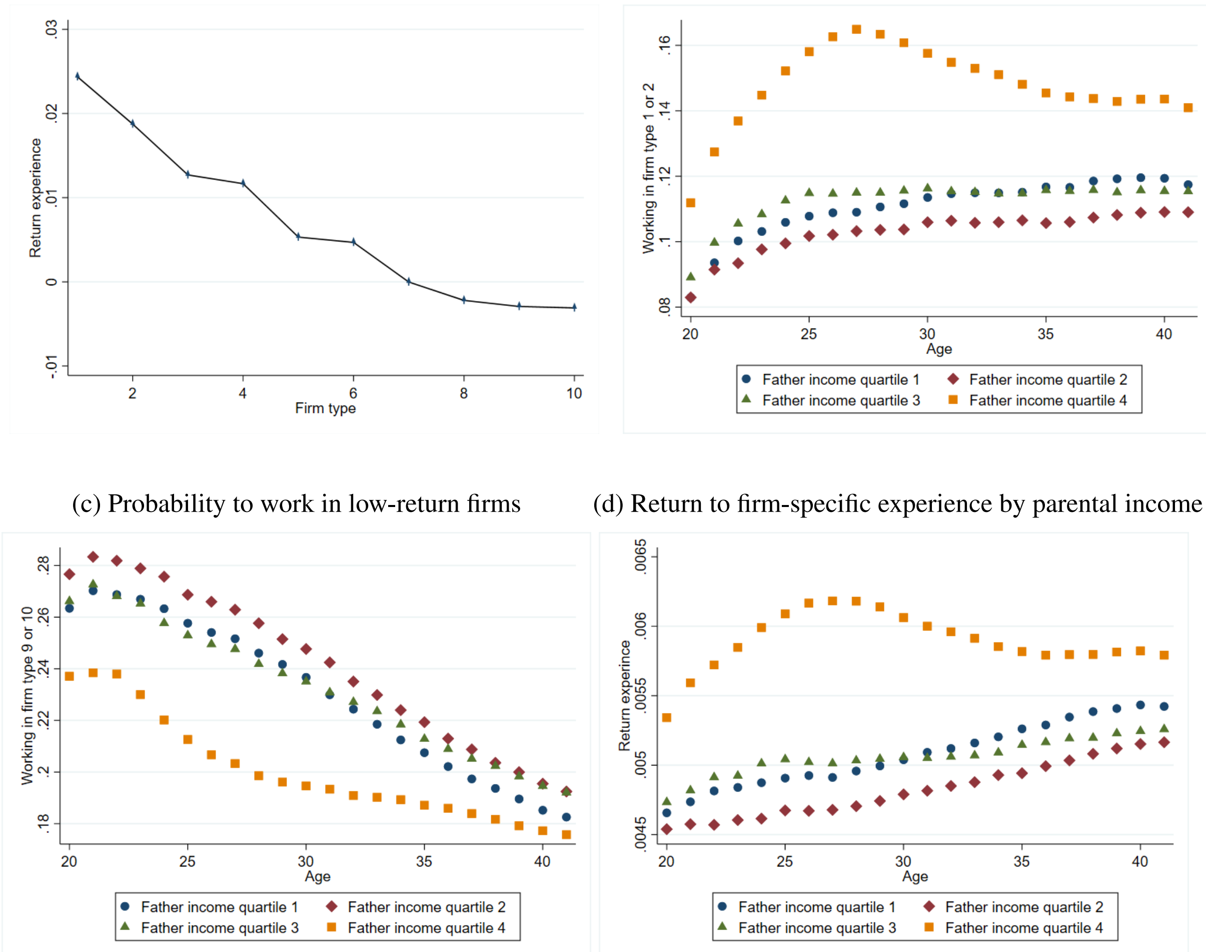
Figure 7: Work in high- vs- low-return firms

(a) Returns to experience in different firm classes

(b) Probability to work in high-return firms

(c) Probability to work in low-return firms

(d) Return to firm-specific experience by parental income



Notes: Sub-figure (a) shows estimates of the coefficients on firm-class experience from equation (3). Sub-figures (b) and (c) show the probability to work in the firm class with the highest returns (firm types 1 or 2) and lowest returns (firm types 9 or 10), by quartile of father's income. Sub-figure (d) shows the average firm-specific returns to experience, by quartile of father's income.

25 (column 3) and the *change* in firm effects between ages 25 and 40 (column 4). Almost half of the SES gap in firm pay at age 40 – and 12% of the IGE – can be attributed to changes in firm pay over age (i.e., climbing the firm ladder, see Section 3.2). Moreover, the sorting of children from affluent families into firms with higher returns to experience explains another 7% of the intergenerational elasticity (column 5).

Since the identification of the return to the firm-specific experience comes from workers who currently are employed in the same firm but have earlier experience from different firm classes, the return component does not capture returns that may crystalize only from moving to other, better-paying firms. Thus, the contribution of returns to firm classes in column 5 of Table 3 can be seen as a lower bound of the contribution of firm-specific returns. Indeed, Figure A12b in Appendix A4 shows that there is indeed a positive relationship between the returns to firm experience γ_m (as captured by column 5) and the change in firm premia (column 4), indicating that firm-specific returns to experience could also explain why children

Table 4: Decomposition of the intergenerational earnings elasticity at age 40

	Dependent variable					
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_j$ at age 25	$\Delta\hat{\psi}_j$ between age 25-40	Returns to firm experience	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}}+\hat{\varepsilon}_{ijt}$
	(1)	(2)	(3)	(4)	(5)	(6)
$y_{f(i)}$	0.178***	0.099***	0.029***	0.021***	0.012***	0.015***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Share of IGE	100%	56%	16%	12%	7%	8%
Worker obs.	284,318	284,318	284,318	284,318	284,318	284,318

Notes: Column (1) reports the estimated slope coefficient from a regression of log child earnings at ages 40 on log father's earnings. Columns (2)-(6) report the slope coefficients from the corresponding regressions when decomposing child log earnings y_{ijt} according to equation (3) into the individual fixed effect α_i, the firm fixed effect at age 25, the change in the firm fixed effect between age 25 and 40, return to firm-specific experience and time-varying controls. The sample differs from our main intergenerational sample since equation (3) is estimated using workers born between 1967-1977 and half of the sample is used to cluster firms into firm classes instead. Moreover, the sample is restricted to individuals who have an observed firm at age 25. Robust standard errors in parentheses.

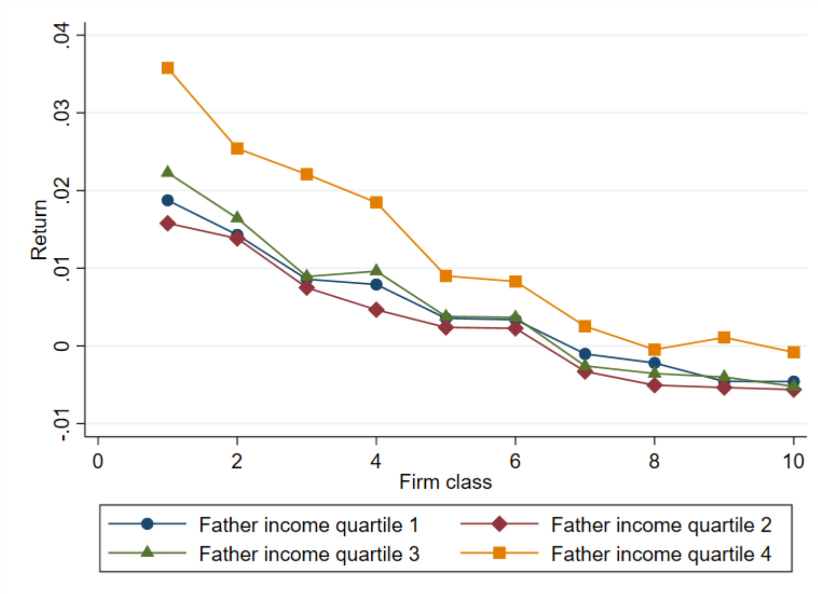
from high-income parents climb the firm ladder faster. Taken together, this dynamic view of firm pay premia suggests an even greater contribution of firms play to intergenerational earnings persistence than the static decomposition in Table 2. Adding columns 3-5 suggests that firms can explain at least 35% of the IGE (or 38% net of column 6).

One worry is that some of these dynamics are driven by the fact that children from high-income families always have higher returns to experience and that they are sorted into particular types of firms. The differences in returns across firm classes as shown in Figure 7a might therefore reflect heterogeneity across individuals rather than firms. Arellano-Bover and Saltiel (2021) test for this concern by interacting firm-specific experience with worker-fixed effects. We follow a similar approach and allow the returns to experience to vary with the father's earnings quartile to estimate the regression:

$$y_{ijt} = \alpha_i + \psi_j + \sum_{m=1}^{K} \gamma_m Exp(m)_{it} + \sum_{m=1}^{K} \delta_m Exp(m)_{it} * \theta_i + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijt}, \tag{4}$$

where θ_i is a vector of dummy variables indicating the quartile of the income of the father. Thus, we allow the returns to experience to vary depending on the father's income. Indeed, Figure 8 shows that children from high-income families have higher returns to experience *within* all firm types. However, all children independent of family background benefit from working in firms with higher returns, and the variation in returns across firm classes is nearly as large as in our estimates based on equation (3) that did not allow for returns to vary with the father's earnings quartile (cf. Figure 7a). In fact, for children from the top quartile of father's earnings, returns vary *more* when allowing for variation by family background.

Figure 8: Returns in different firm classes by father's income



Notes: The figure shows estimates of the coefficients on firm-class experience from equation 4, by quartile of father's income.

4 The role of sorting across firms by skill

We found that some of the SES gradient in firm pay is due to assortative matching between firms and workers. Children from high-income families tend to be more educated and productive (according to the permanent worker component), and prior research shows that workers with higher permanent productivity (or skills) tend to sort into firms that pay higher premia (e.g., Card et al., 2013). But as we also showed, two thirds of the SES gradient in firm pay remains when controlling for the individual fixed effects from the AKM regression (see Table 3, column 4).[15] In the remainder, we refer to this specification as the "conditional firm pay gradient".

There are a two potential interpretations as to why much of the firm pay gradient remains, even conditional on this proxy for individual skills. First and foremost, parental income could indeed have a direct effect on firm sorting beyond what is mechanically driven by skill advantages among their children. Such direct effects could arise due to multiple sources, including informational advantages, social and co-worker networks, and credit or other constraints.[16] They could also arise if preferences for non-monetary amenities differ across families of different income levels, such that part of the gradient in firm pay reflects compensating dif-

[15]Note that we condition on the *estimated* individual fixed effect from the main AKM regression. Thus, we do not include a new set of individual fixed effects in the the regression of estimated firm pay premia on (log) parental income, which would obviously be collinear with parental income.

[16]For example, credit constraints in early age could force poorer children into safe but low-paying jobs (see Staiger, 2022).

ferentials or other non-pay attributes of firms (see Section 5).

However, an alternative interpretation is that the estimated worker effects $\hat{\alpha}_i$ do not capture the full extent of skill sorting, as they are only incomplete measures of skill. First, the estimates $\hat{\alpha}_i$ are only a noisy measure of the α_i component in the AKM model, and such measurement error will bias the estimated contribution of assortative matching to the firm pay gradient (as shown formally by Dobbin and Zohar, 2023). Second, the individual fixed effects α_i capture *all* persistent within-firm differences in earnings, not just those that are due to differences in productivity. For example, persistent taste-based discrimination between ethnic groups would here be falsely interpreted as skill differentials, contributing to skill sorting.[17] A third possibility relates to the multidimensionality of skills. One may surmise that worker-firm sorting is on specific dimensions of skill (e.g., cognitive skills) rather than the entire bundle of skills that contributes to the individual fixed effects in log earnings – and parental background might be more or less strongly associated with those specific sorting dimensions than with other dimensions of skill.

To analyze skill sorting more thoroughly, we consider direct measures on cognitive and non-cognitive skills from military enlistment tests, which we use in addition to our fixed effects-based measures.[18] Specifically, we use a decomposition similar to Gelbach (2016) and Hjorth-Trolle and Landersø (2023) to study sorting on (i) cognitive skills, (ii) social skills, (iii) education and (iv) the estimated individual fixed effect from the AKM model. As the enlistment test was only mandatory for males, we restrict the analyses in this section to males.[19] The skill measures from this test are highly informative about labor productivity, as demonstrated by their strong associations with wages and other long-term labor-market outcomes (Lindqvist and Vestman, 2011; Nybom, 2017). This allows us to directly test the

[17] Moreover, Dobbin and Zohar (2023) note that the AKM worker effects α_i may reflect "social capital", if parents help their children not only to secure a job in better-paying firms, but also to be promoted to better-paying jobs within those firms. While plausible, Stinson and Wignall (2018), San (2022) and Staiger (2022) find that most of the gains from parental networks come from working at a high-wage firm rather than from wage advantages within the firm. And in principle, parental connections might even have a negative effect on wages. For example, Bello and Morchio (2022) predict that "occupational followers" who choose their father's occupation earn lower wages, due to skill mismatch. To address these potential limitations in the interpretability of the AKM fixed effects α_i, we consider here more direct measures of skill.

[18] Dobbin and Zohar (2023) implement two alternative approaches to study the role of assortative matching. First, they develop an instrumental variable approach that uses the child's education as an instrument for their worker fixed effect α_i, which under plausible assumptions provides an upper bound for the contribution of assortative matching to the firm pay gradient. Second, they use education and demographic group as observable proxies for human and social capital, which under alternative assumptions also provides an upper bound for the contribution of assortative matching.

[19] The military tests are taken at around age 18 and were compulsory for all men in the cohorts that we study. The overall cognitive skill score represents an aggregated score from four subtests that measure verbal, logical, spatial and technical skills. The non-cognitive/social test score is based on a half-hour semi-structured interview with a certified psychologist who grades the enlistee along dimensions such as leadership, social maturity, and emotional stability, but also in an overall sense (for further details, see e.g. Lindqvist and Vestman, 2011). We standardize both the overall cognitive and non-cognitive scores to mean zero and standard deviation one, separately for each birth year.

hypothesis that the AKM fixed effects that are typically conditioned on are incomplete measures of skill, such that sorting is underestimated. Moreover, we can compare the relevance of different dimensions of skills (cognitive, social, etc).

The decomposition, summarized by the regression equations (5a)-(7), parses out how much of the relationship between children's firm premia and parents' log income can be attributed to various factors influencing child log income. Having estimated the firm pay gradient β_{firm} in (5a), we then augment this regression with our four mediators of interest: cognitive skill, social skill, education, and the AKM individual fixed effect.

$$\hat{\psi}_j = \mu_\psi + \beta_{firm} y_{f(i)} + \omega_i \tag{5a}$$

$$\hat{\psi}_j = \mu_{\psi,res} + \beta_{firm,res} y_{f(i)} + \beta_{cog} cog_i + \beta_{soc} social_i + \beta_{edu} edu_i + \beta_{akm} \hat{\alpha}_i + \upsilon_i \tag{5b}$$

The coefficient $\beta_{firm,res}$ in this augmented regression (5b) captures the "direct" effect of family background not mediated by skills, while $\beta_{firm} - \beta_{firm,res}$ captures the part explained by the mediators. We then run a set of auxiliary regressions (6a)-(6d) to pin down how closely related each of the mediators are to parental income,

$$cog_i = \mu_{cog} + \phi_{cog} y_{f(i)} + \epsilon_{1i} \tag{6a}$$

$$social_i = \mu_{soc} + \phi_{soc} y_{f(i)} + \epsilon_{2i} \tag{6b}$$

$$edu_i = \mu_{edu} + \phi_{edu} y_{f(i)} + \epsilon_{3i} \tag{6c}$$

$$\hat{\alpha}_i = \mu_{akm} + \phi_{akm} y_{f(i)} + \epsilon_{4i} \tag{6d}$$

The part of the firm pay gradient explained by skills is then given by

$$\beta_{firm} - \beta_{firm,res} = \beta_{cog}\phi_{cog} + \beta_{soc}\phi_{soc} + \beta_{edu}\phi_{edu} + \beta_{akm}\phi_{akm} \tag{7}$$

Table 5 presents the results from this decomposition. We again focus on firm premia observed at age 40. Column (1) replicates the baseline estimate of the relationship between firm pay and (log) parental income, β_{firm}, for the sample of males with observed skill measures. The estimate for this sample of enlisted males (0.056) is very similar to the ones for all males (0.058, see Table A6) and the full population (0.054, see Table 3). Column (2) shows that nearly half of the SES gradient in column (1) is due to sorting on our various skill measures. On the flip side, about 52% of the estimated β_{firm} cannot be explained by skills, suggesting that parental background has a substantial direct effect on firm sorting beyond what is mechanically driven by gaps in the skills of children. Further, columns (3)-(6) show the contribution of each of the respective skill measures to overall sorting (i.e. the components $\beta\phi$ from equation (7)). The individual fixed effects and cognitive skills are both key in explaining why children from richer families work at better-paying firms, with each contributing about

Table 5: The contribution of skills to the SES gradient in firm pay

	Firm pay premia					
	Overall β_{firm}	Unexp. β_{res}	Cognitive $\beta_{cog}\phi_{cog}$	Social $\beta_{soc}\phi_{soc}$	Education $\beta_{edu}\phi_{edu}$	Indiv. FE $\beta_{iFE}\phi_{iFE}$
	0.056***	0.029***	0.011***	0.001***	0.004***	0.012***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Share	100%	51.8%	19.6%	1.8%	7.1%	21.4%
Obs.	371,411	371,411	371,411	371,411	371,411	371,411

Notes: The table reports estimates from the decomposition outlined by equations (5a)-(7) using the main sample, but excluding women and males with missing enlistment scores. Standard errors obtained by 250 bootstraps.

20% to the firm pay gradient β_{firm}. While the education and social/non-cognitive skill components are substantially smaller, they are not inconsequential and together contribute nearly 10%.[20]

Table 5 confirms that the role of skill sorting is underestimated when approximating skill solely by the estimated fixed effects from an AKM model (as conjectured by Dobbin and Zohar, 2023). In Section 2, we reported a conditional firm pay gradient of roughly 0.037, corresponding to an unexplained part of 68.5% (0.037/0.054), increasing to 70.7% when considering males (0.041/0.058, see Appendix Table A6). But when adding our full set of skill measures, the unexplained part drops to 51.8% (0.029/0.056). Cognitive skills are particularly important, explaining 20% of the SES gradient in firm pay, even conditional on estimated individual fixed effects and other controls. However, our analysis also indicate that despite a rich set of skill measures, half of the parental-income gradient in firm pay cannot be explained by skills, thus suggesting that parental income plays a key role for firm sorting beyond what is "mechanically" driven by child skills.

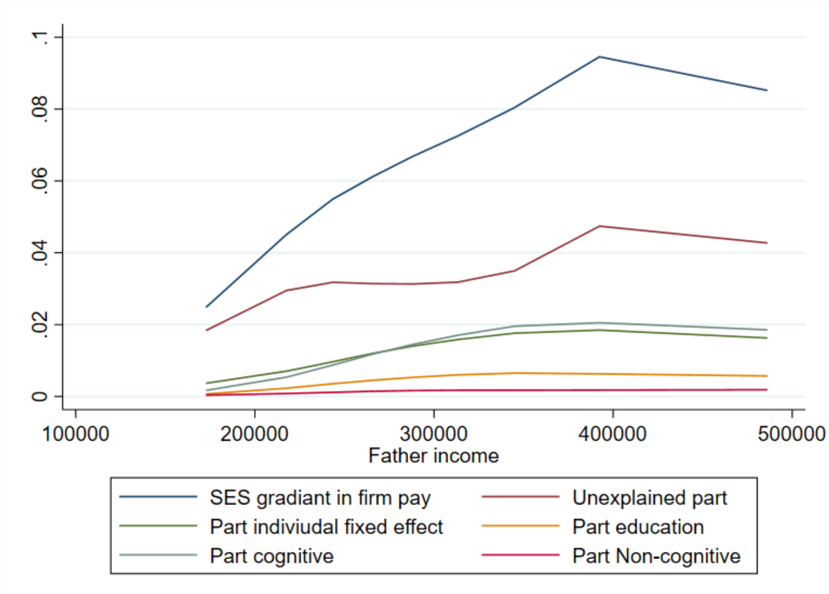
Next, Figure 9 and 10 show the same type of decomposition but across the distribution of parental income and over child age, respectively. To analyze how skill sorting varies with parent income, we follow Hjorth-Trolle and Landersø (2023) and run local-linear regressions of equations (5a)-(7) at each decile of parental income.[21] Figure 9 again demonstrate that the firm gradient generally increases in strength along the distribution, reaching almost 0.1 at high levels of parental income. Moreover, the unexplained part not attributable to sorting grows in absolute size along the distribution but decreases as a share of $\beta_{premium}$. The cognitive skill measure and the individual fixed effect grow steadily in importance, and higher up in the

[20]That education contributes much less to the firm pay gradient is interesting, given that it correlates *more* strongly with parental income than the other mediators (see Appendix Table A7). One possible interpretation is that the worker fixed effects from the AKM model capture differences in formal education better than differences in cognitive skills, but that the latter are an important determinant of worker-firm sorting.

[21]The local linear regression estimates a linear regression around the income mean at each decile of parental income, using a bandwidth of 50 000 Swedish kronor and an epan kernel.

distribution they are important mediators for the raw relationship between firm premia and parental income.

Figure 9: Local linear decomposition of skills along the parent income distribution



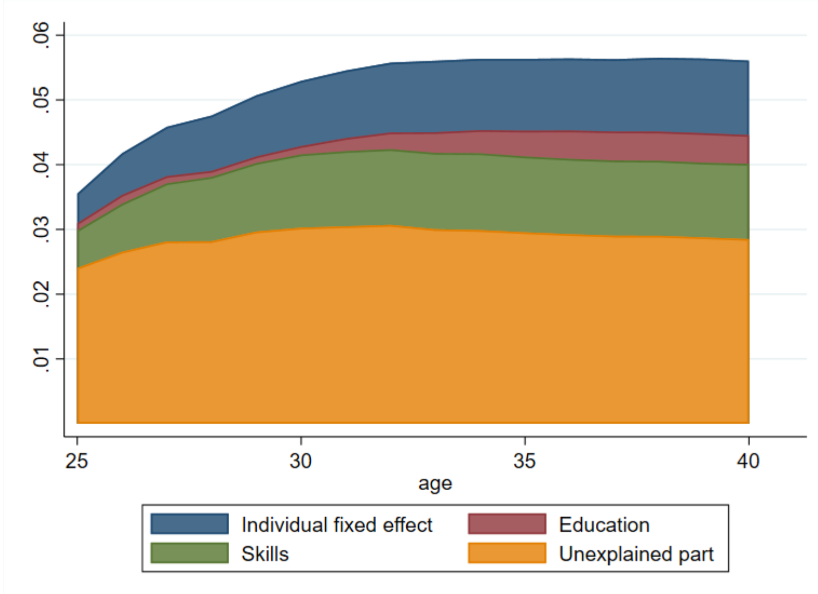
Notes: The figure shows local linear regressions, from the decomposition outlined by equations (5a)-(7) using the main sample, but excluding women and males with missing enlistment scores. The local linear regression estimates a linear regression around the income mean at each decile of parental income, using a bandwidth of 50,000 Swedish kronor and an epan kernel.

Figure 10 shows the decomposition separately for each age over the lifecycle. The part of the firm pay gradient explained by skill sorting, $\beta_{firm} - \beta_{firm,res}$, increases substantially over age. In contrast, the direct effect of family background not mediated by skills, $\beta_{firm,res}$, is already large at age 25 and grows only slightly in size over age. Its relative contribution to the overall firm gradient β_{firm} decreases substantially: nearly 70% of the firm pay gradient at age 25 is due to direct family effects, falling to just 50% at age 40. The finding that family background effects unrelated to skill play a relatively more important role in the early career is intuitive, as children are then likely more closely linked to their parents and parental networks and contacts might be more useful. As children age, skills become more important for firm sorting, contributing further to the overall influence of family background on firm sorting.

5 Do firm pay premia reflect compensating differentials?

Pay is not the only firm attribute that matters to workers, and other aspects of the firm – such as its location, average workloads or fringe benefits – also vary across firms. It is therefore not necessarily the case that high-paying firms are more desirable firms. As these other aspects may also co-vary with parental background, the SES gradient in firm pay may under- or

Figure 10: The contribution of skills to the SES gradient in firm pay by age



Notes: The figure shows age specific estimates from the decomposition outlined by equations (5a)-(7) using the main sample, but excluding women and males with missing enlistment scores.The components sum up to the SES gradient in firm premia.

overstate the role of firms in the intergenerational transmission of advantages. To fix ideas, decompose the firm pay premium as

$$\psi_j = r_j - \kappa a_j \tag{8}$$

where the pay premium of firm j is the sum of a firm-specific worker rent r_j and an amenity component, κa_j. We think of r_j as arising when a firm has some monetary rents to share with its workers, e.g. due to frictions and/or other forms of imperfect competition. The amenity a_j can be either positive, if the firm offers attractive non-pay attributes (e.g. temporal flexibility), or negative, if the firm offers bad non-pay attributes (e.g. poor work environment). Together with the non-negative amenity price $\kappa > 0$, which depends on the marginal worker's preferences, the amenity component, κa_j, thus constitutes a firm-specific compensating differential. We thus assume that a_j is perfectly priced and paid for on the market, and thus will be uncorrelated with worker utility. Consequently, if high-premium firms are overall more desirable it points to the importance of rents, while a weaker association between firm pay premia and the actual attractiveness of firms is evidence of compensating differentials.

Whether the SES gradient in firm pay is primarily due to rents or amenity compensation is crucial for its interpretation, and for the interpreation of intergenerational mobility estimates more generally. Perhaps the SES gradient arises because high-income families place a stronger value on consumption and/or are less averse to bad working conditions. In this scenario, children from high-income families end up in better-paying firms, but those firms are actually worse in other dimensions: measures of intergenerational income persistence would then overstate the extent to which levels of welfare persist across generations. Alternatively, high-income families are better equipped with contacts and networks, information, or other resources, and therefore end up in firms that pay their workers more, conditional on the level of amenities they offer. In this scenario, and if the SES gradient in amenity compensation is small, the gradient in firm pay would approximate the corresponding gradient in welfare. Or maybe firms with high pay tend to be *better* in other dimensions, too, and those non-monetary attributes of firms or jobs are generally better in high-income families. Indeed, given their more favorable financial position, children from high-income families might systematically select into firms that have worse pay but better non-pay attributes, all else equal. In this scenario, intergenerational mobility in underlying welfare would be even lower than income-based estimates suggest.

The decomposition above assumes that amenities are always fully priced into pay, which might not be an accurate description of the world. We can therefore consider an extended decomposition:

$$V_j = \psi_j + \kappa a_j + \kappa b_j = r_j + \kappa b_j \tag{9}$$

where we now focus on the overall value of a firm j, denoted V_j, which depends on the rent-part of the firm pay premium and a second component b_j capturing non-pay characteristics of the firm that are not priced into the worker's pay. Note that if a_j is correctly priced then this part of the pay premium has no influence on the value (or utility) of working for a firm. Thus, the overall value of the firm depends potentially on rents and amenities that are "free of charge" (or at least imperfectly priced).

With some way of inferring the overall value or attractiveness of firms (the V_j), we can address a couple of key questions. First, we can explore whether high-paying firms in general also are more desirable firms. The extent to which higher-paying firms are more desirable firms can then be seen as evidence of rents, while the extent to which this is not the case is evidence of compensating differentials (Sorkin, 2018). In the extreme case, if variation in firm pay is solely due to compensating differentials, there is no relation between the value of firms and their pay premia. Second, we can study whether there is an SES gradient in firm values in a similar fashion as we did for firm pay. By further conditioning on α_i we can infer to what extent the SES gradients in firm value arise from skill sorting or not. Finally, under assumptions of $Corr(a_j, b_j)$, we can estimate SES gradients in firm value conditional on the firm pay premium and infer whether b_j is systematically related to SES.

To explore the different sources to the SES gradient in firm pay, the overall value of firms, and the role of compensating differentials, we use various alternative analyses that all involve using some proxy measure of V_j. Many of the analyses exploit worker transitions between firms, and it thus becomes crucial to distinguish voluntary from involuntary employer-to-employer transitions. Our main strategy here is to focus on voluntary moves, which we define as transitions without any intermediate period of unemployment.[22] First, we explore how parental income relates to alternative measures of a firm's attractiveness that also capture non-pay characteristics, such as the firm's "poaching" and quit rates. We then employ a revealed-preferences based approach similar to Sorkin (2018), which infers the overall values of firms (V_j) from worker transitions between firms.

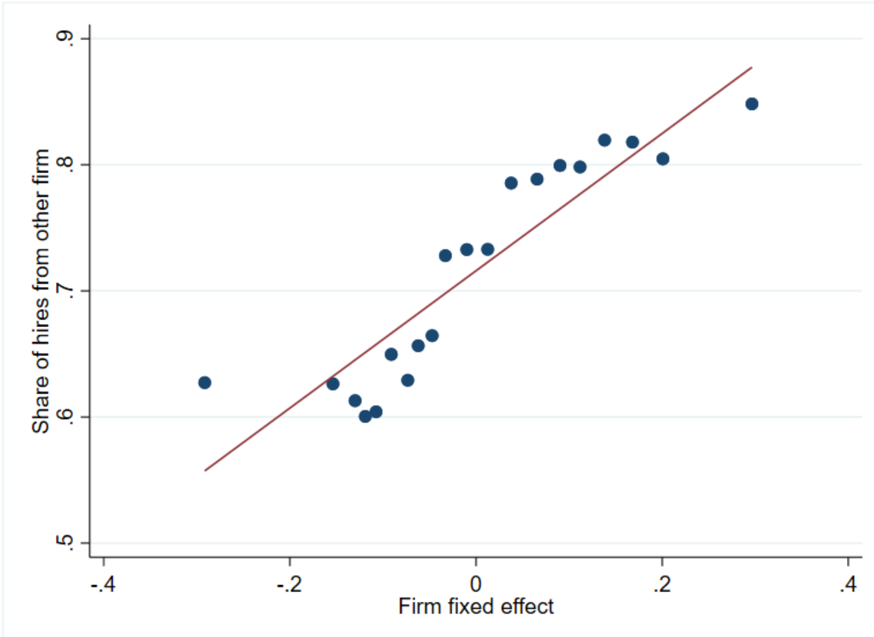
5.1 Alternative measures of the attractiveness of firms

Are high-paying firms indeed more desirable firms? Figure 11 shows that new hires in high-paying firms mostly arrive from employment in other firms, i.e. the new employees have been "poached" from other firms. In contrast, low-paying firms often hire individuals from non-employment, who are less likely to have strong outside options at the time of their hire. This pattern is consistent with the "job ladder" from standard search models (Burdett and Mortensen, 1998), and indicates that high-paying firms are indeed more attractive from the perspective of workers.[23] Given our simple decomposition above, the result would be inconsistent with that most or all of the firm pay premium reflects compensating differentials (priced non-pay characteristics).

[22]We identify unemployment periods using data on UI benefit receipts, and include in our analyses only employer-to-employer transitions associated with zero received benefits.

[23]However, while this relationship may hold on average, it does not necessarily follow that it also holds for the SES-gradient in firm pay (i.e., the way children from high-SES background select into firms may differ from the average relations observed in the labor market). In the next section, we quantify the attractiveness of each firm, to then study this question in more details.

Figure 11: Poaching rate from other firms



Notes: Binned scatterplot of the share of hires from employment ("poaching rate") on the firm fixed effects estimated based on the AKM model in equation (1).

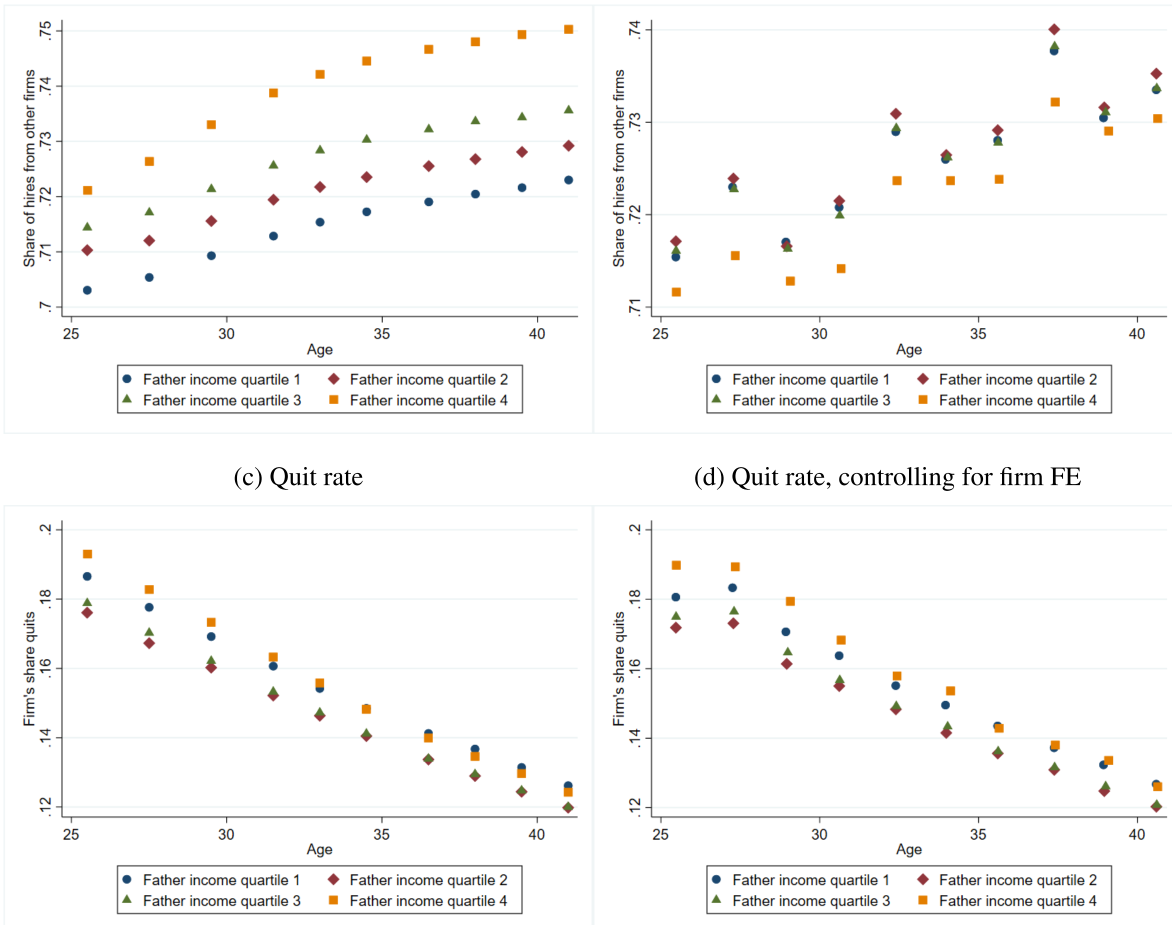
In Figure 12a, we study whether the firms' poaching rate varies systematically by parental background. Indeed, we see a similar pattern by parental income in the poaching rate as we do in firm pay premia (see Figure 2a): the gaps open up already at early age, and widen further up to the mid 30s. However, Figure 12b shows that *conditional* on firm pay, high-SES children end up in firms with slightly *lower* poaching rates. This might indicate that those firms are not as desirable as they seem in terms of firm pay, although the gaps in the conditional poaching rate in Figure 12b are much smaller than the unconditional gaps in Figure 12a. Overall, we find that high-SES children do sort into more attractive firms (as proxied by firm poaching rates).

In Figures 12c and 12d we show the corresponding pattern in the *quit rates* of firms. Again, the idea is that firms that for monetary or non-monetary reasons are attractive employers will see fewer quits among their existing workforce, as it is harder for other firms to poach workers from these firms. The evidence here is more mixed: early in their careers, children from the bottom but also from the top SES quartile work in firms that have slightly *higher* quit rates. The quit rates generally fall with age, and so do the gaps between SES groups. The patterns are also less sensitive to the inclusion of firm fixed effects.

Figure 12: Poaching and quit rates

(a) Poaching rate from other firms (b) Poaching rate, controlling for firm FE

(c) Quit rate (d) Quit rate, controlling for firm FE



Notes: The figure shows "poaching" and quit rates over the lifecycle by father's income quartile. Sub-figure (a) shows the mean poaching rate, defined as the share of hires from employment. Sub-figure (b) shows the same poaching rate but controlling for the estimated firm premium. Sub-figure (c) shows the quit rate at the firm, and sub-figure (d) shows quit rates controlling for the estimated firm premium.

5.2 Inferring SES gradients in firm values using revealed preferences

While the previous analyses of poaching and quit rates provided insights, a concern is that these measures are imperfect proxies of the overall values of firms. We therefore employ a more comprehensive way of inferring firm values, or V_j from above, using the revealed-preference measure of firm values based on worker flows across firms suggested by Sorkin (2018). One can think of Sorkin's approach as the Google PageRank algorithm, but for firms. In short, the idea is that if workers voluntarily move from one firm to another it must imply that the value of the destination firm is higher. We define voluntary employer-to-employer transitions as transitions where workers receive no unemployment benefits or have a year of zero earnings, in between adjacent employment spells associated with different firms. We thus retrieve a value for each firm, which could consist of both rents and (non-priced) amenities. While showing that children from high-SES families work in firms with higher

Table 6: Firm values and compensating differentials

	Dependent variable				
	$\hat{\psi}_{j=J(i,t)}$	$\hat{V}_j$	$\hat{V}_j$	$\hat{V}_j$	$\hat{V}_j$
	(1)	(2)	(3)	(4)	(5)
$y_{f(i)}$	0.063***	0.034***	0.016***	0.016***	0.006
	(0.001)	(0.004)	(0.004)	(0.004)	(0.004)
$\hat{\psi}_{j=J(i,t)}$				0.286***	0.233***
				(0.012)	(0.013)
$\hat{\alpha}_i$			0.148***		0.110***
			(0.007)		(0.007)
Observations	356,774	356,774	356,774	356,774	356,774

Notes: Column (1) reports the slope coefficient from regressing $\hat{\psi}_j$ from equation (1) on father's log income for the subsample of firms that are included in the model to estimate the firm values. Column (2) reports the slope coefficient from regressing the estimated firm value, $\hat{V}_j$, following Sorkin (2018), on father log income. Columns (3)-(5) show slope coefficient estimates from regressing firm values on father log income including different controls. Column (3) controls for $\hat{\alpha}_i$, column (4) controls for $\hat{\psi}_j$, and column (5) controls for $\hat{\psi}_j$ and $\hat{\alpha}_i$ simultaneously. Robust standard errors in parentheses.

pay premia is not necessarily proof of them being better off in a welfare sense, studying the same gradients in terms of firm values allows us to draw inference about welfare differences. Table 6 shows results for how firm values relate to father's log income.

Because we cannot retrieve estimates of V_j for all firms – they need to be a part of a more restrictive *strongly* connected set in terms of voluntary firm-to-firm transitions (see Sorkin, 2018) – the sample size is reduced considerably. In column 1, we thus re-estimate the SES gradient in firm pay premia, which is slightly larger in this more restricted sample than in our baseline (.063 vs. .057). In column 2, we then document that the SES gradient extends to the estimated firm value, and thus the overall desirability of the firms. Thus, it seems like sorting across employers actually make children from high-SES families better off. When controlling for the individual fixed effect (column 3), the SES gradient is substantially weakened but remains positive and significant. Thus, one reason why high-SES children are able to enter more desirable firms is that they have higher skills, which enable them to sort into higher value firms. It is notable that skill sorting appears more important for firm values than for firm pay premia, which is intuitive if people use their skills to maximize their overall welfare (rather than just their pay check).[24] However, even conditional on skills, high-SES children are considerably more likely to end up in higher value firms.

The firm value consists of both rents and non-priced amenities, and without further assumptions we cannot distinguish the relative roles of these components. The results are also silent about whether high-SES children experience positive or negative amenities — it only says that the value of the firm premium is larger than any potential negative amenities. Thus,

[24]Controlling for the individual fixed effect diminishes the SES gradient in firm value by 53% but the firm pay premium "only" by 31% (see Table 2).

and together with the fact that firm pay premia and firm values are positively correlated, there is evidence that not all of the variation in pay premia is due to compensating differentials – there is some room for rents.

In column 4, we show the SES gradient in firm value conditional on the firm premium, thus comparing children from different backgrounds that work in firms with similar pay premia. In doing so, the SES gradient in terms of firm value decreases in size but remains positive. Thus, when working in firms with similar pay premia, people from high-SES families are able to enter firms with relatively higher rents and non-priced amenities than priced (bad) amenities compared to low-SES children. However, without further assumptions we cannot know if this is since high-SES children earn higher rents or higher non-priced amenities given a certain firm premium. Obviously, we might assume away non-priced amenities, and define the firm value as a proportional function of firm-specific rents. In that case, the estimates in column 4 would suggest an SES gradient in rents, even conditional on firm pay premia. However, when we in addition control for the individual fixed effect (column 5), the SES gradient becomes insignificant, indicating that the main reason that high-SES children are able to enter high-value firms, given the firm premium, is skill sorting.

6 Conclusions

This paper examined the extent to which the sorting of workers across firms contributes to intergenerational earnings persistence. We build on the large literature on the drivers of intergenerational persistence. While the literature has traditionally focused on childhood development and inequalities in parental investments in their children's human capital, we add by providing a labor-market perspective. In particular, we use Swedish administrative data and decompose earnings into permanent individual components (approximating productivity) and firm-specific pay premia a la Abowd et al. (1999) and many others in their footsteps. We then add data enabling us to link parents to children, and provide a multitude of evidence on the SES gradient in the firm portion of pay, how it evolves over the lifecycle, its underlying drivers, and how the gradient ought to be interpreted.

Our findings indicate that disparities in firm pay premia can account for a significant portion of the intergenerational elasticity of income in Sweden. This suggests that the advantages or disadvantages associated with people's family backgrounds can have lasting impacts on their career trajectories and long-run outcomes in life. The emergence of SES gaps in firm pay already at the outset of one's career implies that individuals from more privileged backgrounds have access to more favorable entry points into the labor market. These advantages are compounded by the fact that they are able to climb the firm pay ladder faster, frequently switching employers and securing higher pay gains conditional on such changes. While skill sorting – the fact that high-SES children tend to have higher skills, and highly skilled people sort into better firms – accounts for a sizable portion of the widening of these pay gaps, a large share of the SES gradient in firm pay remains also conditional on a very detailed set of controls for skill. Furthermore, our results remain robust even after accounting for compensating differentials and alternative measures of firm quality. Thus, high-SES children sort into firms that not only deliver larger pay checks, but ultimately also higher overall welfare.

References

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

AGRAWAL, D. R., E. J. JAHN, AND E. JANEBA (2024): "Do Commuting Subsidies Drive Workers to Better Firms," Tech. rep., LASER Discussion Paper No. 149.

ANDREWS, M. J., L. GILL, T. SCHANK, AND R. UPWARD (2008): "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171, 673–697.

ARELLANO-BOVER, J. AND F. SALTIEL (2021): "Differences in On-the-Job Learning Across Firms," .

BECKER, G. S. AND N. TOMES (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 1153–1189.

BELLO, S. L. AND I. MORCHIO (2022): "Like Father, Like Son: Occupational Choice, Intergenerational Persistence and Misallocation," *Quantitative Economics*, 13, 629–679.

BINGLEY, P., M. CORAK, AND N. WESTERGÅRD-NIELSEN (2012): "The Intergenerational Transmission of Employers in Canada and Denmark," in *From Parents to Children: The Intergenerational Transmission of Advantage*, ed. by J. Ermisch, M. Jantti, and T. M. Smeeding, New York: Russell Sage.

BJÖRKLUND, A., M. LINDAHL, AND E. PLUG (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121, 999–1029.

BJÖRKLUND, A., J. ROINE, AND D. WALDENSTRÖM (2012): "Intergenerational Top Income mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?" *Journal of Public Economics*, 96, 474–484.

BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2005): "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 95, 437–449.

BONHOMME, S., K. HOLZHEU, T. LAMADON, E. MANRESA, M. MOGSTAD, AND B. SETZLER (2023): "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*, 41, 291–322.

BRANDÉN, G. AND M. NYBOM (2019): "Utvecklingen av intergenerationell rörlighet i Sverige," Tech. rep., Swedish Ministry of Finance.

BURDETT, K. AND D. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–73.

CARD, D., A. R. CARDOSO, J. HEINING, AND P. KLINE (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36, S13 – S70.

CARD, D., A. R. CARDOSO, AND P. KLINE (2016): "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly journal of economics*, 131, 633–686.

CARD, D., J. HEINING, AND P. KLINE (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 128, 967–1015.

CORAK, M. AND P. PIRAINO (2010): "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 29, 37–68.

CORNELISSEN, T., C. DUSTMANN, AND U. SCHÖNBERG (2017): "Peer effects in the workplace," *American Economic Review*, 107, 425–56.

DEUTSCHER, N. AND B. MAZUMDER (2023): "Measuring Intergenerational Income Mobility: A Synthesis of Approaches," *Journal of Economic Literature*, 61, 988–1036.

DOBBIN, C. AND T. ZOHAR (2023): "Quantifying the Role of Firms in Intergenerational Mobility," Unpublished manuscript.

EECKHOUT, J. (2018): "Sorting in the Labor Market," *Annual Review of Economics*, 10, 1–29.

ENGBOM, N., C. MOSER, AND J. SAUERMANN (2023): "Firm pay dynamics," *Journal of Econometrics*, 233, 396–423.

ENGZELL, P. AND C. MOOD (2023): "Understanding Patterns and Trends in Income Mobility through Multiverse Analysis," *American Sociological Review*, forthcoming.

ENGZELL, P. AND N. WILMERS (2024): "Firms and the Intergenerational Transmission of Labor Market Advantage," .

GELBACH, J. B. (2016): "When do covariates matter? And which ones, and how much?" *Journal of Labor Economics*, 34, 509–543.

GERARD, F., L. LAGOS, E. SEVERNINI, AND D. CARD (2021): "Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil," *American Economic Review*, 111, 3418–57.

GRÖNQVIST, E., B. ÖCKERT, AND J. VLACHOS (2017): "The Intergenerational Transmission of Cognitive and Noncognitive Abilities," *Journal of Human Resources*, 52, 887–918.

HECKMAN, J. J. AND S. MOSSO (2014): "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 6, 689–733.

HJORTH-TROLLE, A. AND R. LANDERSØ (2023): "The Different Sources of Intergenerational Income Mobility in High and Low Income Families," Tech. rep., Rockwool Foundation Research Report Study Paper 195.

JÄNTTI, M. AND S. JENKINS (2014): "Income Mobility," in *Handbook of Income Distribution*, ed. by A. Atkinson and F. Bourguignon, Elsevier, vol. 2, 807–935.

KLINE, P., R. SAGGIO, AND M. SØLVSTEN (2020): "Leave-out estimation of variance components," *Econometrica*, 88, 1859–1898.

KRAMARZ, F. AND O. N. SKANS (2014): "When Strong Ties are Strong: Networks and Youth Labour Market Entry," *The Review of Economic Studies*, 81, 1164–1200.

KREMER, M. (1993): "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 108, 551–75.

LACHOWSKA, M., A. MAS, AND S. A. WOODBURY (2020): "Sources of displaced workers' long-term earnings losses," *American Economic Review*, 110, 3231–66.

LE BARBANCHON, T., R. RATHELOT, AND A. ROULET (2020): "Gender Differences in Job Search: Trading off Commute against Wage*," *The Quarterly Journal of Economics*, 136, 381–426.

LEE, S. Y. AND A. SESHADRI (2019): "On the intergenerational transmission of economic status," *Journal of Political Economy*, 127, 855–921.

LINDQVIST, E. AND R. VESTMAN (2011): "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 3, 101–128.

LOCHNER, L. AND A. MONGE-NARANJO (2012): "Credit Constraints in Education," *Annual Review of Economics*, 4, 225–256.

MANNING, A. (2013): "Monopsony in motion," in *Monopsony in Motion*, Princeton University Press.

NYBOM, M. (2017): "The distribution of lifetime earnings returns to college," *Journal of Labor Economics*, 35, 903–952.

NYBOM, M. AND J. STUHLER (2016): “Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation,” *Journal of Human Resources*, 51.

SAN, S. (2022): “Who Works Where and Why: The Role of Social Connections in the Labor Market,” Tech. rep., Working Paper, Hebrew University of Jerusalem.

SOLON, G. (1992): “Intergenerational Income Mobility in the United States,” *American Economic Review*, 82, 393–408.

SONG, J., D. J. PRICE, F. GUVENEN, N. BLOOM, AND T. VON WACHTER (2019): “Firming up inequality,” *The Quarterly Journal of Economics*, 134, 1–50.

SORKIN, I. (2018): “Ranking firms using revealed preference,” *The Quarterly Journal of Economics*, 133, 1331–1393.

STAIGER, M. (2022): “The Intergenerational Transmission of Employers and the Earnings of Young Workers,” Tech. rep., Working Paper, Harvard University.

STINSON, M. AND C. WIGNALL (2018): “Fathers, Children, and the Intergenerational Transmission of Employers,” Working Papers 18-12, Center for Economic Studies, U.S. Census Bureau.

ZIMMERMAN, S. D. (2019): “Elite colleges and upward mobility to top jobs and top incomes,” *American Economic Review*, 109, 1–47.

Online Appendix

A1 Additional results

A1.1 Variance decomposition of the AKM model

Table A1: Variance decomposition

	AKM sample	Main sample (born 1967-77)	
Age	20-64	25-41	39-41
	(1)	(2)	(3)
Variance of log earnings	0.273	0.248	0.208
Components:			
Individual FEs	0.103 (37.7%)	0.073 (29.4%)	0.074(35.6%)
Firm FEs	0.020 (7.3%)	0.023 (9.3%)	0.023(11.1%)
Covariance (sorting)	0.019 (7.0%)	0.021 (8.5%)	0.023 (11.1)
Covariates and residual	0.131 (48.0%)	0.129 (52.0%)	0.088(42.3%)
Worker obs.	7,668,377	967,417	857,064
Number of firms	341,798	228,285	118,258
Worker-year obs.	126,475,937	13,550,074	2,437,567

Notes: The table shows a variance decomposition of log earnings into the of components of equation (1). Column (1) show the variance decomposition for the AKM sample, column (2) shows the result for the main lifecycle sample and column (3) shows the result for the main sample with mean earnings estimated for the ages 39-41.

Using our estimates from equation (1), we can decompose the variance in income as

$$\begin{aligned} Var(y_{ijt}) &= Var(\alpha_i) + Var(\psi_j) + 2Cov(\alpha_i, \psi_j) + Var(\boldsymbol{X}_{it}\boldsymbol{\delta}) \\ &\quad +2Cov(\boldsymbol{X}_{it}\boldsymbol{\delta}, \alpha_i + \psi_j) + Var(\varepsilon_{ijt}) \end{aligned} \tag{A1}$$

We report the results in Appendix Table A1, separately for three samples: our AKM sample, our main intergenerational sample across the entire age span (ages 25-41), and our main sample at age 39-41. The first two terms on the right-hand side in equation (A1) describe what fraction of the overall earnings variance is due to individual and firm components, respectively. The third component measures the contribution of worker-firm sorting; if this covariance is positive, there is positive assortative matching in the sense that workers with high (permanent) unobserved productivity sort into firms with high pay premia. The last three terms capture earnings variation due to covariates and the error term.

As found by others, the most important component for explaining the variance of log earnings is the (variance of) worker effects, here at 29-38% across the three samples. On the other hand, firm fixed effects and the covariance between firm and worker fixed effects together explain 14-22% of the total variance. When we compare the samples, we see that the variance decomposition is largely stable across samples. However, for the main sample observed over the lifecycle (column 2) we find a somewhat decreased importance of the individual component, compared to the full AKM sample (column 1). For the prime-age version of the main sample (column 3), which only includes incomes at ages 39-41, there is a slight uptick in the importance of firms and sorting (rows 2 and 3) compared to the baseline. Overall, the decomposition is very similar to Engbom et al. (2023) who use similar data and specifications, and also largely in line with evidence from the US (e.g. Song et al., 2019).

A1.2 Decomposition of the IGE

A1.2.1 Measurement error

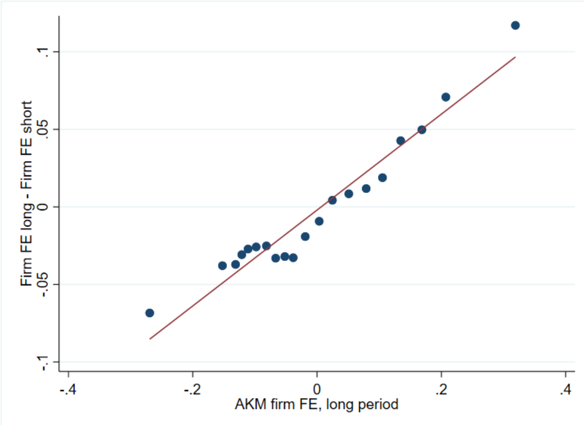
The table shows the IGE decomposition when the AKM model have been estimated for shorter time period, for the years 2010-2015. Mean earnings and mean firm fixed effect are then calculated for all these years.

Table A2: Decomposition of the IGE, AKM estimated for the years 2010-2015

	Dependent variable			
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_{j=J(i,t)}$	$\hat{\psi}_{j=J(i,t)}$
	(1)	(2)	(3)	(4)
A: AKM estimated for the years 2010-2015				
$y_{f(i)}$	0.197***	0.165***	0.031***	0.021***
	(0.001)	(0.001)	(0.000)	(0.000)
$\hat{\alpha}_i$				0.059***
				(0.000)
Share of IGE	1.0	0.84	0.16	0.11
Worker obs.	784,259	784,259	784,259	784,259
B: AKM estimated for 1985-2018 (intergenerational sample same as in panel A)				
$y_{f(i)}$	0.197***	0.117***	0.051***	0.033***
	(0.001)	(0.001)	(0.000)	(0.000)
$\hat{\alpha}_i$				0.154***
				(0.001)
Share of IGE	1.0	0.59	0.26	0.17
Worker obs	784,259	784,259	784,259	784,259

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings over the ages 39-41 on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child log earnings y_{ijt} according to equation (1) into individual fixed effects α_i and mean firm fixed effects ψ_j over the ages 39-41.The columns for time-varying control is not included since we have imputed firm values for the ages 39-4, even if this ages are outside of the 2010-2015 window if the firm existed for years 2010-2015, and thus we miss data on the time varying controls for those observations. . In panel A equation (1) is estimated for the years 2010-2015. In panel B equation (1) is estimated for the full time-period 1985-2018, but the observations in the intergenerational regression are limited to the same as in panel A. Robust standard errors in parentheses.

Figure A1: Difference AKM 2010-2015 and long AKM



Notes: The figure shows that difference in the estimated firm premia between the AKM estimated for the long period 1985-2018, and the AKM estimated for the short period 2010-2015 against the firm premia from the AKM estimated for the long period.

Table A3: Decomposition of IGE, random sub-sample

	Dependent variable			
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_{j=J(i,t)}$	$\hat{\psi}_{j=J(i,t)}$
	(1)	(2)	(3)	(4)
A: Dropping 50% of individuals before AKM estimation				
$y_{f(i)}$	0.203***	0.120***	0.053***	0.036***
	(0.002)	(0.001)	(0.001)	(0.001)
$\hat{\alpha}_i$				0.145***
				(0.001)
Share of IGE	1.0	0.59	0.26	0.18
Worker obs.	418,980	418,980	418,980	418,980
B: AKM estimated for full sample (intergenerational sample same as in panel A)				
$y_{f(i)}$	0.203***	0.120***	0.053***	0.035***
	(0.002)	(0.001)	(0.001)	(0.000)
$\hat{\alpha}_i$				0.152***
				(0.001)
Share of IGE	1.0	0.59	0.26	0.17
Worker obs	418,980	418,980	418,980	418,980

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings over the ages 39-41 on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child log earnings y_{ijt} according to equation (1) into individual fixed effects α_i, mean firm fixed effects ψ_j over the ages 39-41, and time-varying controls. In panel A equation (1) is estimated after dropping a random subsample of 50% of all individuals. In panel B equation (1) is estimated for the full sample, but the observations in the intergenerational regression are limited to the same as in panel A. Robust standard errors in parentheses.

A1.2.2 Alternative estimation of the AKM regression

Table A4: Decomposition of the IGE: Alternative estimations of the AKM equation

	Dependent variable				
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_{j=J(i,t)}$	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$	$\hat{\psi}_{j=J(i,t)}\vert\hat{\alpha}_i$
	(1)	(2)	(3)	(4)	(5)
	A. AKM estimated with time-varying firm FE				
$y_{f(i)}$	0.200***	0.125***	0.057***	0.018***	0.040***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.63	0.29	0.09	0.20
Worker obs.	847,447	847,447	847,447	847,447	847,447
	B. AKM estimated using establishment codes				
$y_{f(i)}$	0.198***	0.105***	0.072***	0.022***	0.056***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.53	0.36	0.11	0.28
Worker obs.	831,927	831,927	831,927	831,927	831,927
	C. AKM estimated using establishment codes for large firms				
$y_{f(i)}$	0.198***	0.107***	0.069***	0.023***	0.052***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.54	0.35	0.12	0.26
Worker obs.	836,554	836,554	836,554	836,554	836,554
	D. AKM estimated without excluding firms with few movers				
$y_{f(i)}$	0.199***	0.115***	0.058***	0.027***	0.043***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.58	0.29	0.14	0.22
Worker obs.	904,384	904,384	904,384	904,384	904,384

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings at ages 39-41on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child earnings y_{ijt} according to equation (1) into individual fixed effects α_i, mean of firm fixed effects ψ_j for ages 39-41, and time-varying controls. The different panel shows different variants of estimating equation (1). In panel A we estimate time-varying firm fixed effects by dividing the period 1985-2018 into 4 periods, and allow the firm-fixed effects to vary between the periods. In panel B we use establishment codes instead of firm codes to estimate equation (1). In Panel C establishment codes are used for large firms and firm codes for small firms with 1,000 or fewer unique workers during the analysis period. In panel D we estimate the AKM without excluding firms with few movers.

Table A5: Decomposition of the IGE: AKM estimated with wages

	Dependent variable				
	y_{ijt} (1)	$\hat{\alpha}_i$ (2)	$\hat{\psi}_{j=J(i,t)}$ (3)	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$ (4)	$\hat{\psi}_{j=J(i,t)}$ (5)
	A: AKM estimated with wages				
$y_{f(i)}$	0.175***	0.132***	0.022***	0.021***	0.008***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
$\hat{\alpha}_i$					0.105***
					(0.001)
Share of IGE	1.0	0.75	0.13	0.12	0.05
Worker obs.	565,231	565,231	565,231	565,231	565,231
	B: AKM estimated with earnings (excluding individuals not in wage sample)				
$y_{f(i)}$	0.194***	0.120***	0.044***	0.030***	0.025***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
$\hat{\alpha}_i$					0.156***
					(0.001)
Share of IGE	1.0	0.62	0.23	0.15	0.13
Worker obs.	565,231	565,231	565,231	565,231	565,231

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child wages at ages 39-41 on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child wage y_{ijt} according to equation (1) into individual fixed effects α_i, mean of firm fixed effects ψ_j for ages 39-41, and time-varying controls. In Panel A we estimate equations (1) and (2) using the wage structure sample, which covers roughly a third of private sector employees (with those in larger firms oversampled) and all public sector employees, in total corresponding to about 50% of the workforce. In Panel B we estimate equations (1) and (2) using earnings, but limiting the observations to the same observations as in the wage sample. Robust standard errors in parentheses.

A1.2.3 Decomposition of the IGE: Heterogeneity

Table A6 shows the decomposition of the intergenerational earnings elasticity for different subsamples.

Table A6: Decomposition of the IGE: Heterogeneity

	Dependent variable				
	y_{ijt} (1)	$\hat{\alpha}_i$ (2)	$\hat{\psi}_{j=J(i,t)}$ (3)	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}}+\hat{\varepsilon}_{ijt}$ (4)	$\hat{\psi}_{j=J(i,t)}\vert\hat{\alpha}_i$ (5)
	A. Sample: Men				
$y_{f(i)}$	0.230***	0.135***	0.058***	0.036***	0.041***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Share of IGE	1.0	0.59	0.25	0.16	0.18
Worker obs.	436,709	436,709	436,709	436,709	436,709
	B. Sample: Women				
$y_{f(i)}$	0.169***	0.097***	0.050***	0.022***	0.040***
	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.57	0.30	0.13	0.24
Worker obs.	420,355	420,355	420,355	420,355	420,355
	C. Excluding workers who work in same firm as father				
$y_{f(i)}$	0.191***	0.112***	0.051***	0.028***	0.033***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.59	0.27	0.15	0.17
Worker obs.	748,293	748,293	748,293	748,293	748,293
	D. Excluding public sector firms				
$y_{f(i)}$	0.228***	0.124***	0.068***	0.035***	0.051***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.54	0.30	0.15	0.22
Worker obs.	549,635	549,635	549,635	549,635	549,635
	E. Excluding firms with less then 10 movers				
$y_{f(i)}$	0.200***	0.118***	0.052***	0.030***	0.034***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.59	0.26	0.15	0.17
Worker obs.	836,199	836,199	836,199	836,199	836,199
	F. Excluding firms with less then 50 movers				
$y_{f(i)}$	0.201***	0.121***	0.047***	0.032***	0.028***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	1.0	0.60	0.23	0.16	0.14
Worker obs.	745,473	745,473	745,473	745,473	745,473

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings at ages 39-41 on log father's earnings. Columns (2)-(4) report the slope coefficients from the corresponding regressions when decomposing child earnings y_{ijt} according to equation (1) into individual fixed effects α_i, mean of firm fixed effects ψ_j for ages 39-41, and time-varying controls. Panel A shows results for males, and panel B shows results for women. In Panel C, workers who work in the same firm as their fathers are excluded, where working in the same firms as the father is defined as having ever worked in the father's main firm (main firm is the firm the father works in for most years between 1985-2018). Panel D shows results where public sector firms are excluded, where public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99. Panel E excludes firms with less than 10 movers and Panel F excludes firms with less than 50 movers (baseline: at least five movers).

A2 Skill sorting

In this Appendix we provide additional evidence on the decomposition of the SES gradient in firm pay into skill-based sorting ("assortative matching") and residual sorting. Table A7 reports estimates from the auxiliary regressions (6a)-(6d) of each skill measure on parental income (i..e, estimates of $\phi_{cog}, \phi_{soc}, \phi_{edu}$ and ϕ_{akm}). Note that the regression coefficients are not directly comparable, as they also reflect differences in the scaling of each variable. We therefore focus on the correlation coefficient, which is equal to the square root of the R-squared reported in the table. We find that parental income correlates most strongly with child education, while the correlation with the child's social skills is lowest. Despite correlating strongly with parental income, child education contributes only a small share to the firm pay gradient (see Table (5)). The comparatively low correlation for our proxy of social skills is possibly explained by measurement error, as social skill measures tend to be more noisy than measures of cognitive skills (Grönqvist et al., 2017).

Table A8 reports estimates regression (7), showing that the coefficients change only marginally when excluding parental income from the regression.

Table A7: Skills and father's income

	Dependent variable			
	Cognitive skills	Social skills	Education	$\hat{\alpha}_i$
$y_{f(i)}$	1.022***	0.714***	1.477***	0.129***
	(0.007)	(0.006)	(0.009)	(0.001)
Observations	371,411	371,411	371,411	371,411
R-squared	0.060	0.038	0.070	0.052

Notes: The figure shows result from the Gelbach decomposition, equations (6a)-(7), regressing each of the skills on father's income.

Table A8: Firm pay premia and skills

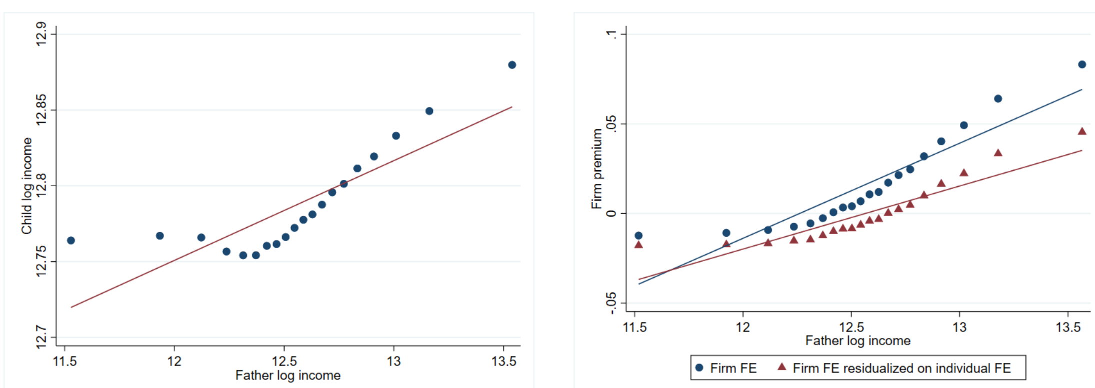
	Dependent variable	
	$\hat{\psi}_{j=J(i,t)}$	$\hat{\psi}_{j=J(i,t)}$
$y_{f(i)}$	0.029***	
	(0.001)	
Cognitive skills	0.010***	0.011***
	(0.000)	(0.000)
Social skills	0.002***	0.002***
	(0.000)	(0.000)
Education	0.003***	0.004***
	(0.000)	(0.000)
$\hat{\alpha}_i$	0.089***	0.096***
	(0.001)	(0.001)
Observations	371,411	371,411
R-squared	0.086	0.080

Notes: The figure shows result from the Gelbach decomposition, equation (5b), regressing the firm premia on father's income and each of the skills.

A2.1 Non-linear firm pay gradients

Figure A2: Child income and firm premia by father's income (logs)

(a) Child vs father log income (b) Firm premium at age 40 vs father's log income

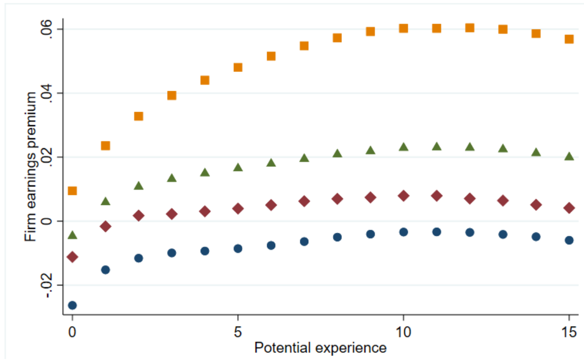


Notes: Sub-figure (a) shows binned scatter plots of child's log income at age 40 by father's log income. Sub-figure (b) shows firm fixed effects ψ_j at age 40 estimated by equation (1) and firm premia residualized on individual fixed effect by father log income.

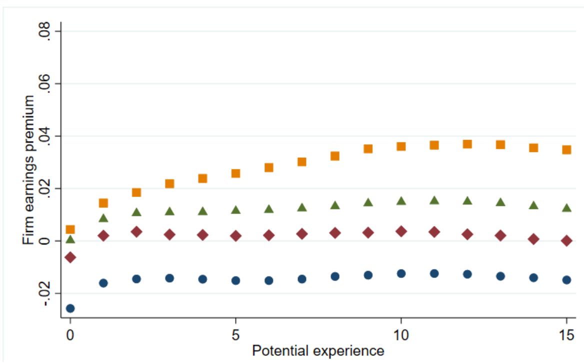
A3 Additional evidence on lifecycle dynamics

Figure A3: Firm earnings premium over the lifecycle using potential experience

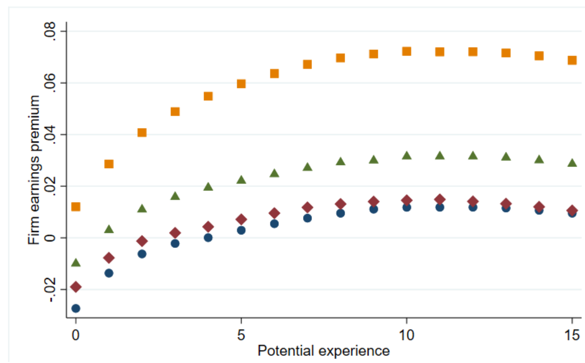
(a) Full sample

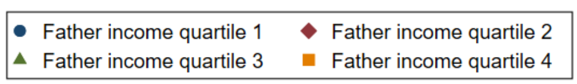


(b) Non-college



(c) College





Notes: The figures plots the mean estimated firm premium $\hat{\psi}_j$ against potential experience over the life cycle, by quartile of father's income. Potential experience is defined as the year minus the the year the individual enters the labor market. Entering the labor market is defined as the first year, after age 20 of having higher then low earnings (where low earning is defined as 20% of the median earnings of men aged 45), and after the age for potential finishing school (defined as age- (years of education age - education +6)). Sub-figure (a) shows the result for the full sample, sub-figure (b) shows the result for children without collage education and sub-figure (c) shows the result for children with college education. Father's income quartiles are defined in the full sample.

Figure A4: Firm pay premium over the lifecycle conditional on individual fixed effects

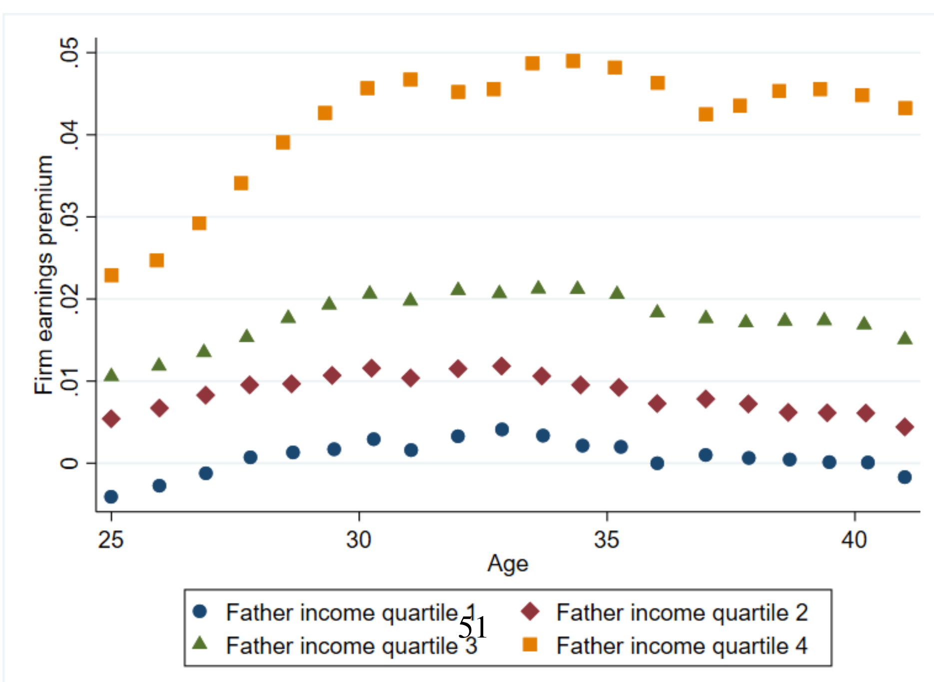
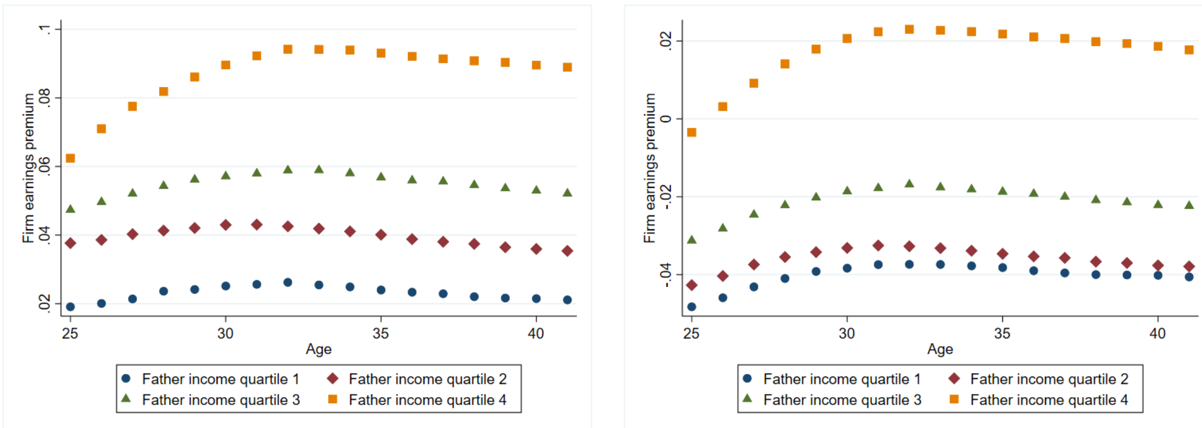


Figure A5: Firm premia over the life-cycle by gender

(a) Men (b) Women

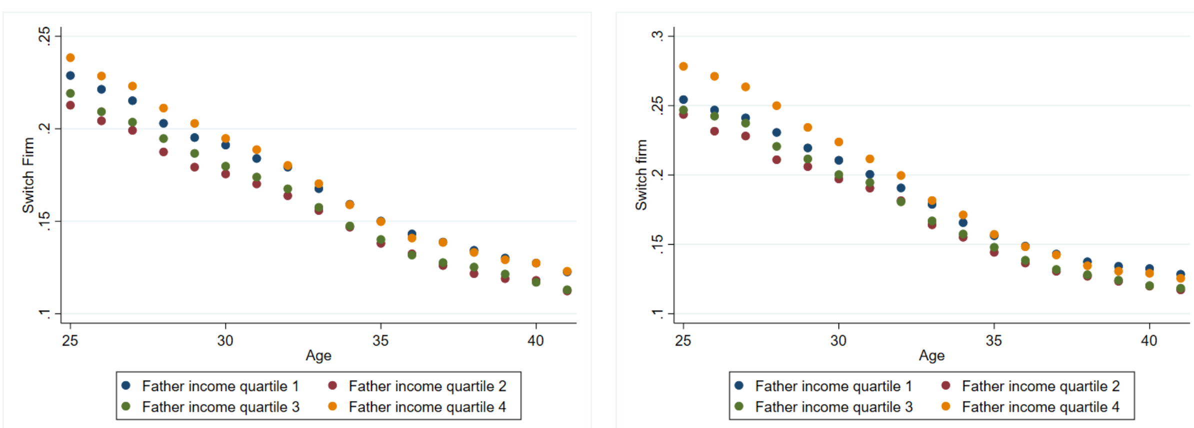


Notes: The figures show the estimated firm premium $\hat{\psi}_j$ from equation (1) over the life cycle, by quartile of father's income. Sub-figure (a) shows the results for men and (b) shows the result for women. Father's income quartiles are defined in the full sample.

A3.1 Firm switching

Figure A6: Firm switches by education

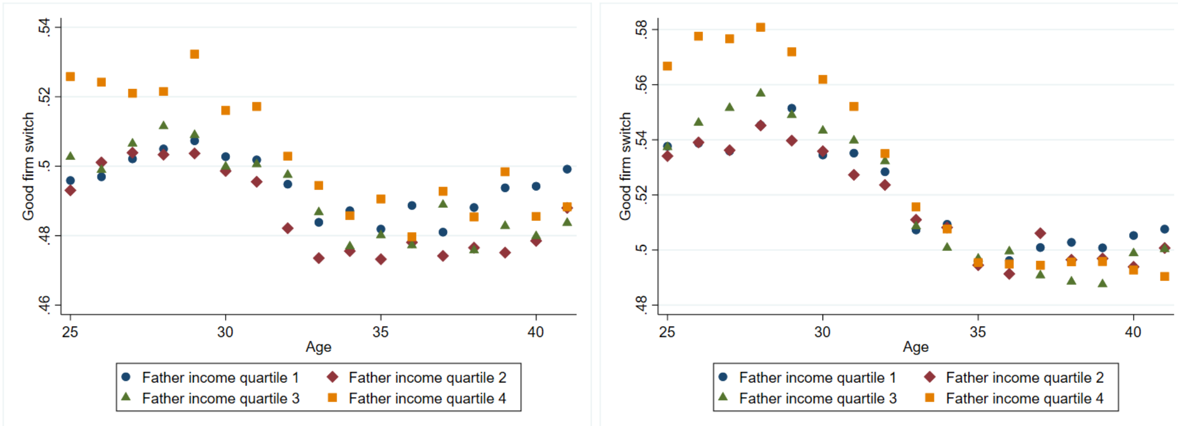
(a) Non-college (b) College



Notes: The figures show pattern for switching firms over the life-cyle by quartile of father's income. Sub-figure (a) shows the result for individuals who do not have a college education and sub-figure (b) shows the results for n individuals who have a college education.

Figure A7: Proportion of premium-improving firm switches

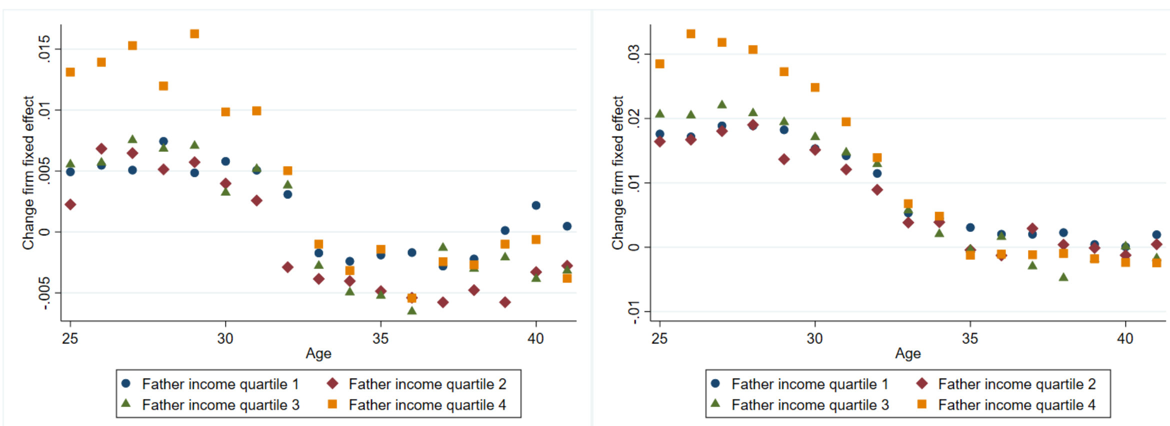
(a) Non-college (b) College



Note: The figure shows the probability of switching to a firm with a higher firm premium than the one before conditional on switching, by father's income quartile. Sub-figure (a) shows the result for individuals who do not have a college education and sub-figure (b) shows the results for individuals who have a college education.

Figure A8: Mean change in firm premium among switchers

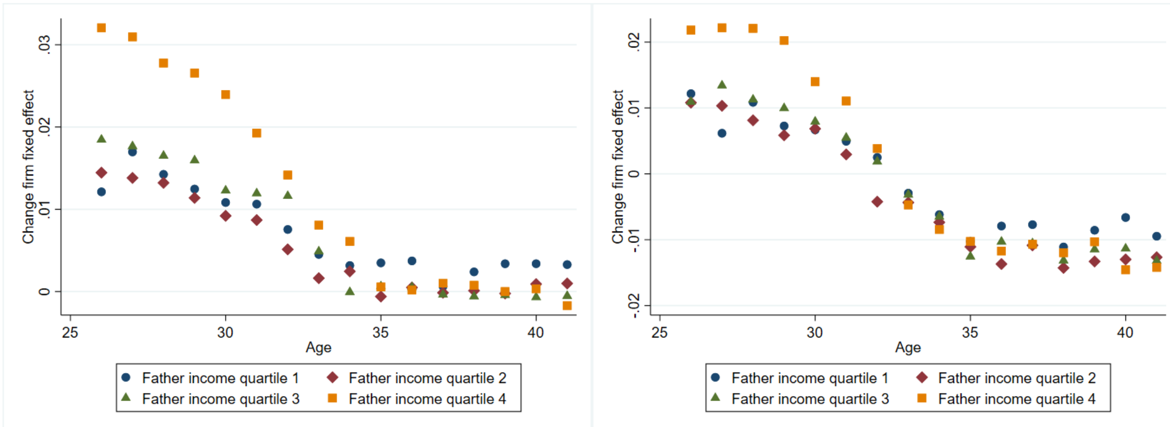
(a) Non-college (b) College



Note: The figure shows the difference between the new firm premia and the firm premia before for individuals who switch firms. Sub-figure (a) shows the result for individuals who do not have a college education and sub-figure (b) shows the results for individuals who have a college education.

Figure A9: Change in firm FE for voluntary and involuntary switches

(a) Voluntary switches only (b) Involuntary switches only



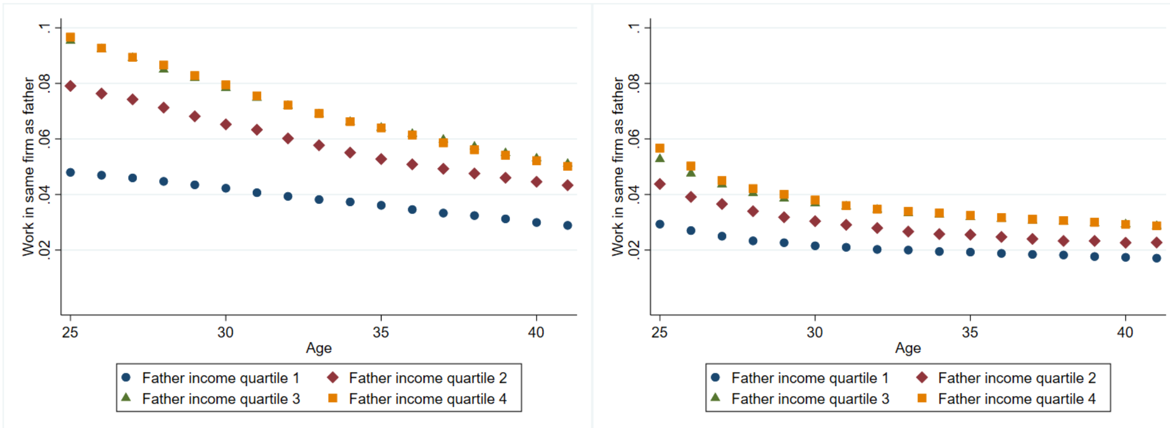
Note: The figure shows the difference between the new firm premia and the firm premia before for individuals who switch firms, by father's income quartile. Sub-figure (a) shows the results for voluntary switches, defined as a switch without any unemployment insurance or without any year with zero income. Sub-figure (b) shows the result for involuntary switches, defined as a switch with either unemployment insurance or a year of zero income between working at the old firm and starting at the new firm.

A3.2 Working patterns

In this section, we provide additional evidence on how working and commuting patterns vary by parental background. Figure A10 plots the share of individuals who work in the same firm as their father, separately for those with and without a college degree. Figure A10 plots the share of individuals who commute (i.e., work and reside in different municipalities), separately for those with and without a college degree. Children from high-income parents are more likely to commute, even within education group. Finally, Table A9 provides event-study type regression results conditional on individual fixed effects, showing how firm pay changes for individuals who begin to commute. Also when we consider such within-individual variation, the firm premium increases when individuals start to commute. As shown in column 1, commuting raises firm premia by about 1 pp. The pay benefit of commuting grows somewhat with age (column 2). Finally, column 3 shows that individuals from different parental backgrounds in terms of father's income benefit similarly from commuting. Thus, we conclude that firm pay is positively related to commuting and in a similar manner irrespective of parental background, but that high-SES children are more likely to commute.

Figure A10: Working in the same firm as father

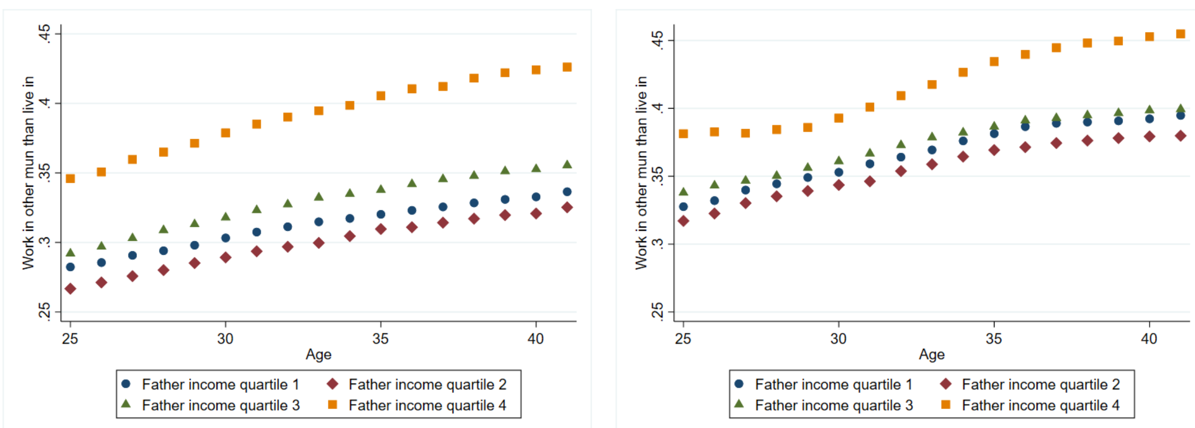
(a) Non-college (b) College



Note: The figure shows the proportion of children who work in the father's main firm at different ages, separately by quartile of fathers' income. The father's main firm is defined as the firm the father works in for the most years between the years 1985-2018. Sub-figure (a) include children without college education and sub-figure (b) include children with college education.

Figure A11: Commuting

(a) Non-college (b) College



Notes: The figure shows the proportion of individuals who commute (i.e., work in another municipality than they live in) over the life-cycle, by quartile of father's income. Sub-figure (a) includes children without college education and sub-figure (b) children with college education.

Table A9: Commuting

	Dependent variable: $\hat{\psi}_{j=J(i,t)}$		
	(1)	(2)	(3)
Commuting indicator	0.010***	0.017***	0.011***
	(0.000)	(0.000)	(0.000)
Commuting # 2nd father's income quartile			-0.001**
			(0.000)
Commuting # 3rd father's income quartile			0.000
			(0.000)
Commuting # 4th father's income quartile			-0.000
			(0.000)
Commuting # age normalized at 40		0.001***	
		(0.000)	
Age controls	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

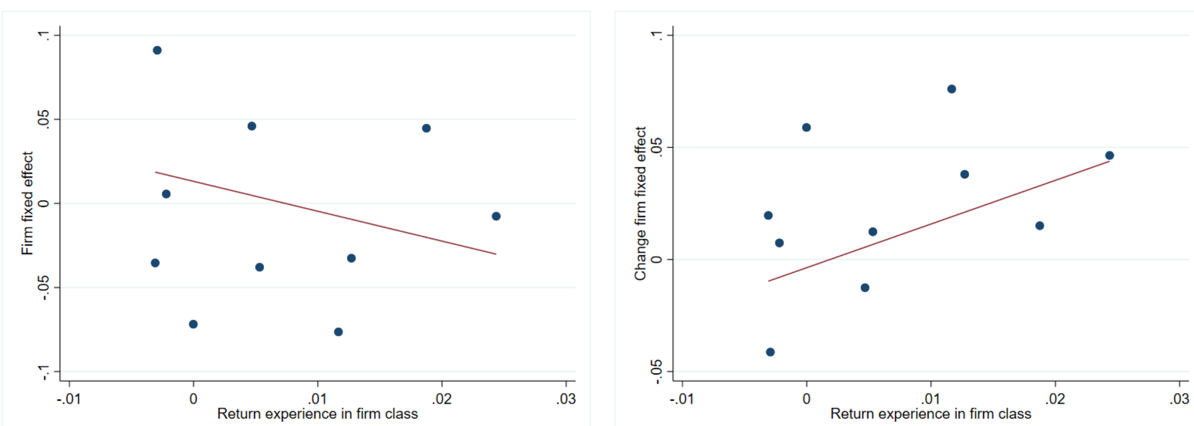
Notes:: The table shows results with the firm premia as the dependent variable. The regressions include a dummy variable for working in a municipality other than living in and individual fixed effects. Columns 1 and 3 include flexible age controls (age, age squared, age interaction with father income quartile, and age squared interacted with father income quartile). Column (2) includes linear age dummies normalized at age 40 to ease interpretation and includes dummies for working in other municipalities than living in interacted with age. Column (3), includes dummies for working in other municipalities than living in interacted with the father income quartile.

A4 Return to experience in firm classes

Figure A12: Firm fixed effects and return to experience

(a) Firm fixed effects and return to experience

(b) Change in firm fixed effects and return to experience



Notes: The figures show firm fixed effects and experience estimated with regression (3). Sub-figure (a) shows the relationship between firm fixed effects and returns to experience in the firm class. Sub-figure (b) shows the relationship between the change in firm fixed effects for individuals who change firms, and the return to experience in the firm class the individuals switch from.