# Labor-market Drivers of Intergenerational Earnings Persistence

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

January 11, 2026

**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 36% of the intergenerational elasticity of earnings in Sweden, rising to 50% 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, the Research Network for Intergenerational Mobility, EALE, and University of Bergen for helpful comments.

# Introduction

Children from high-income families earn substantially more as adults than children from low-income families (Jäntti and Jenkins 2014; Deutscher and Mazumder, 2023; Nybom, 2024). A central question in the social sciences is why this is the case. Education is viewed as a key mediator (Goldthorpe, 2014), and seminal contributions in economics 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 literature decomposes intergenerational persistence into pre- and post-birth factors to disentangle nature and nurture (e.g. Björklund et al., 2006; Black et al., 2005). As such, much of the literature focuses on how differences in human capital accumulation *in childhood* translate into pay gaps in adulthood. In contrast, few studies focus on the role of labor-market mechanisms in shaping intergenerational income persistence.

In this paper, we study how the sorting of workers across firms contributes to intergenerational persistence. We first show that worker-firm sorting is a key factor why earnings differ with parental socioeconomic status (SES): differences in firm *pay premia* account for 36% of the intergenerational elasticity of earnings in Sweden, rising to 50% when including dynamic returns to firm-specific experience. Two thirds of the static pay gaps open already at career start, suggesting that children from more privileged backgrounds find more favorable entry points to the labor market. Moreover, their pay advantages widen in early career as they climb the firm pay ladder faster, switch firms more frequently, and secure higher pay gains conditional on switching. A key mechanism contributing to these patterns is skill sorting, as more productive workers tend to sort into higher-paying firms. However, while explaining the widening of pay gaps over the early career, and around half of the firm pay gradient at mid-age, skill sorting cannot explain why large pay gaps open already at career start. Finally, we consider the role of non-wage firm amenities and show that children from high-income families work not only in better-paying but also more desirable firms.

We start by presenting a set of empirical facts on the socioeconomic gradient in firm pay in Sweden, following a similar approach as a recent study by Dobbin and Zohar (2023). Using population-wide earnings data, combined with employer-employee and parent-child links, we first decompose earnings into worker and firm components using the two-way fixed effects "AKM" framework of Abowd et al. (1999). We then show that the sorting of high-SES workers to better-paying firms explains a substantial part of the intergenerational transmission of income advantages, accounting for more than one third (36%) of the intergenerational earnings elasticity (IGE). Differences in firm pay across regions and selection into higher-paying industries and larger firms account for more than half of this firm pay gradient. 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, contributing to the elevated persistence at the top of Sweden's income distribution (Björklund et al., 2012).

In the paper's second part, we study how the firm pay gradient evolves over the lifecycle. Two-thirds of the firm pay gap opens already at career start, consistent with the notion that children from richer families find more favorable "entry points" to the labor market. Their pay advantage then widens further in early career before stabilizing in the mid-30s. Differences in early-career progression therefore explain part of the firm pay gap, as high-SES children switch firms more frequently and secure larger pay gains conditional on switching. These results can be interpreted through the family of models in which search frictions generate variation in firm pay and "job ladders" (Burdett and Mortensen, 1998; Manning, 2013), suggesting that high-SES children climb the job ladder more quickly. All these patterns hold within education groups but are especially strong among college graduates.

Moreover, firm pay gradients may exist not only in a static sense ("firm fixed effects") but also dynamically, as children from privileged backgrounds sort into expanding firms and those characterized by higher wage growth. To study the implication of such dynamic firm advantages, we follow Arellano-Bover and Saltiel (2021) and Battiston et al. (2024) and use a $k$-means clustering approach to estimate an extended two-way fixed effects model that allows for firm-class specific returns to experience. We show that high-SES children sort into firms with substantially higher wage growth; by age 40, these dynamic firm advantages accumulate to 21% of their overall firm pay advantage, increasing the total (static and dynamic) contribution of firms to 50% of the IGE. Labor market sorting thus explains a large share of the intergenerational persistence in income.

In the third part, we investigate whether this overrepresentation of high-SES children in better-paying firms reflects skill sorting. It is well established that worker and firm effects from the AKM framework correlate positively, as more productive workers sort into higher-paying firms (Card et al., 2013). Given that children from high-income families have on average better skills, the firm pay gradient may simply reflect this skill-based assortative matching. In this case, firms' contribution to SES gaps would be indirect, by amplifying the effect of skill differences and inequities that already arose in childhood. To illustrate the importance of skill sorting, we show that conditioning on the worker fixed effects from the AKM model reduces the SES gradient in firm pay by about 30% – a similar reduction as Dobbin and Zohar (2023) find for Israel, using the same approach.

We add to this evidence in two ways. First, we can test for skill sorting more directly using late-adolescence skill measures from mandatory (for men) military enlistment tests. Conditioning on cognitive and non-cognitive skills, in addition to education and AKM worker effects, reduces the firm pay gradient by about 50%, that is, half of the SES gaps in firm pay at age 40 is due to skill sorting. The cognitive skill measure provides the most additional explanatory power. Second, we can trace how the relative importance of skill sorting and family background changes over the lifecycle. 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 reflects "direct" family effects unrelated to observed skills, while skill sorting explains only 30%. However, skill sorting becomes increasingly important over age, explaining nearly half of the firm pay gradient at mid-age.

The large effect of parental background at labor market entry is consistent with a growing literature showing that parents (Kramarz and Skans, 2014; San, 2022; Staiger, 2022; Eliason et al., 2023) and peers (Zimmerman, 2019; Eliason et al., 2023; Barrios-Fernández et al., 2024; Campa, 2025) facilitate access to higher-paying firms. However, the observation that skill sorting contributes to the firm pay gradient – and that this contribution magnifies over age – illustrates that social networks are not the sole mechanism by which family background affects firm pay. Motivated by the literature on employer learning (Farber and Gibbons, 1996; Altonji and Pierret, 2001), we formulate a simple model of skill sorting under imperfect information to show that statistical discrimination could account for our key empirical findings: why the firm pay gradient opens up already at career start, why both skills and family background net of skills contribute to this gradient, why skill sorting strengthens over the lifecycle, and why the overall firm pay gradient also steepens over age.

In the final part of our paper, we ask whether the SES gradient in firm pay reflects compensating differentials or other non-pay attributes of firms. If intergenerational income transmission partly stems from inherited preferences – for example, if some families place greater weight on income and consumption compared to non-monetary attributes – then measures of income persistence would overstate persistence in welfare. On the other hand, if children from more affluent families sort into firms with more favorable non-pay attributes, in addition to higher pay premia, then intergenerational persistence in welfare might be higher than income-based measures suggest. To our knowledge, there exists very little evidence on these questions.[1]

To assess the role of non-pay attributes of firms, we first explore how firm premia and parental income relate to a widely used proxy for overall firm desirability, the extent to which firms are able to poach workers from other firms. The results confirm that higher-paying firms are more attractive employers, 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) that exploits firm-to-firm transitions of workers to infer firms' non-wage values. Overall, we find little evidence that compensating differentials vary systematically with family background. Instead, skills and other drivers of labor-market advan-

[1] As recent exceptions, Haeck and Laliberté (2025) show that children from high-SES families are overrepresented in occupations characterized by more desirable characteristics, and Schüle (2026) finds that they place less value on income and job security than their less well-off peers.

tages appear more important for intergenerational transmission than correlated preferences for non-wage job attributes.

Our work contributes to a nascent literature on the contribution of labor market factors to intergenerational persistence. A key motivation is Dobbin and Zohar (2023), whose methodological approach we adopt in our baseline analysis. We extend their work by tracing the evolution of pay gaps over the lifecycle and by capturing the dynamic implications of worker-firm sorting. In addition, we use direct skill measures to study skill sorting over the lifecycle and examine the role of non-wage attributes in overall firm desirability. Closely related are Engzell and Wilmers (2025), who take a sociological perspective grounded in stratification theory to study the transmission of firm pay advantages in Sweden, and Laliberté and Whalley (2025), who analyze the role of social connections in firm pay advantages in Canada. Engzell and Wilmers decompose both parent and child earnings into worker and firm components, study the sources of firm premia, and test whether firms pay different premia to workers from different socioeconomic backgrounds. Laliberté and Whalley distinguish the role of connections from skills by conditioning on fine-grained educational information, comparing individuals with the same degree, from the same institution.

In comparison, we find a greater contribution of firms to intergenerational persistence than other studies, explaining 36% of the IGE at mid-age in the static model and even more in the dynamic model. Using a similar specification, Dobbin and Zohar (2023) find that firm sorting explains 23% of the IGE in Israel. Engzell and Wilmers (2025) show that it explains 23% of a rank-based persistence measure in Sweden, while Laliberté and Whalley (2025) report that it explains 30% of the rank-rank slope in Canada. This contrast reflects not just cross-country differences, but also differences in data quality, our use of establishment identifiers for large firms, and – in the dynamic model – the incorporation of firm-specific returns to experience. While finding a greater overall role of firms, we find a similar relative contribution of skill sorting as Dobbin and Zohar (2023), despite using a different methodological approach.

Our work also relates to a growing literature on social connections in worker-firm sorting. In a pioneering study on labor market factors in intergenerational transmission, Corak and Piraino (2010) show that many Canadian children work for the same employer as their parents, especially among high-income families; Bingley et al. (2012) and Stinson and Wignall (2018) confirm these findings for Denmark and the U.S. More generally, Kramarz and Skans (2014), San (2022), Staiger (2022) and Eliason et al. (2023) show that parental and family connections to workplaces ("strong ties") matter for early-career job finding and earnings, especially for lower-educated workers. A related literature emphasizes the importance of peer networks formed in higher education.[2] While these studies isolate the specific effect

[2] Zimmerman (2019), Michelman et al. (2022), and Barrios-Fernández et al. (2024) show that admission to elite colleges improves labor market outcomes primarily by altering social environments rather than academic achievement. Campa (2025) shows that exposure to more affluent peers in school also raises firm pay.

of family or peer connections, we quantify the overall role of worker-firm sorting in intergenerational persistence, which depends also on other mechanisms – such as skill sorting, statistical discrimination, or the transmission of preferences. Our finding that large SES gaps in firm pay emerge already at labor market entry – and that these gaps are not well explained by individual skill – is consistent with evidence on the importance of social connections in early career outcomes. However, we show that statistical discrimination under imperfect information about worker skill can also contribute to these initial pay gaps, helping to understand their magnitude. Moreover, our finding that skill sorting explains most of the *increase* in the firm pay gradient over age suggests that family connections matter primarily at labor market entry rather than for subsequent career progression.

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 far more sorting by parental background than one might expect from skill sorting alone. We further document firm sorting by cognitive and non-cognitive skill (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 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 focus and approach differ, we find a similarly important role of skill sorting for SES pay gradients in Sweden. A recent literature highlight the contribution of labor market sorting to cross-sectional inequality (7-12% in our data, comparable to estimates for other countries reported in Bonhomme et al., 2023). Our findings suggest, however, that differences in firm pay account for a considerably higher share of the *intergenerational* persistence in earnings.

# 1 Data and specifications

## 1.1 Data

Our analyses use longitudinal administrative data from Statistics Sweden linking earnings, employers, education, and family background for parents and children over several decades. We combine tax registers on earnings; full-population employer–employee registers identifying firms and establishments; the education register for highest attained education; and the multigenerational register to link children to parents.

Earnings data cover the full working-age population from 1968 to 2018. Our main outcome is gross annual labor earnings, including self-employment income, bonuses, fringe benefits, and short-term employer-provided sickness benefits. We exclude very low annual earnings observations, to ensure that our estimates are not overly influenced by variation in labor supply.[3] Firm and establishment characteristics, including location and industry codes, are available from 1985 onward. The birth and family registers allow us to link nearly all Swedish-born children born in 1932 or later, and foreign-born children born in 1961 or later, to both parents.

*Main intergenerational sample.* Our primary sample includes men and women born between 1967 and 1977, observed in the labor market from ages 25 to 41, with at least one observed firm fixed effect. These cohorts allow us to observe a substantial portion of individuals' careers while measuring fathers' prime-age income. Fathers' long-run income is measured between ages 45 and 55 using either mean log earnings or mean earnings rank. Fathers' annual earnings are residualized with respect to year dummies and quadratic age effects, and ranks are computed relative to fathers in the main sample. Individuals with missing father income – primarily due to migration – are excluded. In some analyses, we focus on peak career outcomes measured at ages 39-41.

*AKM sample.* To estimate firm premia with an AKM model, we use a matched employer–employee dataset covering the full Swedish labor force ages 20-64 from 1985 to 2018. This dataset includes annual earnings, firm or establishment identifiers, age, gender, and education. Firms are the unit of analysis, except for large firms, for which establishment identifiers are used.[4] Estimated firm and worker fixed effects from this AKM model are inputs into the intergenerational analysis.

*Descriptive statistics.* Table 1 reports summary statistics for both samples. Differences between them mainly reflect age and cohort composition, as well as the inclusion of earlier earnings years in the AKM sample. Individuals in the AKM sample were born, on average, 12 years earlier, resulting in lower college attainment and slightly lower log earnings. The average log firm size is 4.3, corresponding to roughly 80 employees. Applying sample restrictions yields a final analysis sample of 836,743 individuals observed at ages 39-41, and 960,925 individuals observed at some point between ages 25 and 41.[5]

[3]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.

[4]We use establishment identifiers for multi-establishment firms with more than 1,000 employees (on average over the period). We show below that this choice has only limited effects on our results.

[5]Our full-population data include 1,301,551 individuals with positive earnings at ages 39–41. Requiring non-missing links to fathers and observed paternal earnings reduces the sample to 1,076,969 and 933,647 individuals, respectively. Excluding very low earnings for either children or fathers leaves 910,665 observations. Imposing valid firm links with identified firm fixed effects and observed demographics (such as education) yields a final analysis sample of 836,743 individuals at ages 39–41. Extending the sample to individuals observed at any point between ages 25 and 41 increases the full population to 1,441,540 individuals with positive earnings in at least one year, and applying the same restrictions results in a final analysis sample of 960,925 individuals.

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 | 12.48 | 12.55 | 12.80 |
| Share women | 0.49 | 0.48 | 0.48 |
| Share with college degree | 0.39 | 0.48 | 0.50 |
| Year of birth | 1961 | 1972 | 1972 |
| Log firm size | 4.31 | 4.26 | 4.29 |
| Firm size | 455.89 | 435.90 | 460.56 |
| Number individuals | 7,482,149 | 960,925 | 836,743 |
| Number firms | 460,875 | 314,598 | 172,225 |

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, "AKM") to decompose earnings into firm and individual components, conditional on a set of time-varying controls. Specifically, we model the log earnings $y_{ijt}$ of individual $i$ employed in firm $j = J(i,t)$ in year $t$ as:

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

where $\alpha_i$ denotes a worker fixed effect, $\psi_j$ a firm fixed effect (the "firm pay premium"), $\boldsymbol{X}_{it}$ a vector of time-varying controls with coefficient vector $\boldsymbol{\delta}$, and $\varepsilon_{ijt}$ an error term. The controls in $\boldsymbol{X}_{it}$ include year dummies and education-gender-specific age dummies.[6] Because age, cohort (captured by the worker fixed effect), and time are collinear, unrestricted age dummies are not identified. Rather than imposing a parametric functional form on lifecycle earnings profiles, we follow Engbom et al. (2023) and impose age effects to be constant over ages 45–54, where earnings profiles are relatively flat.

The observation that employers offer systematically different wages to the same worker has received much attention in the recent literature. These "firm pay premia" may arise naturally from search frictions (Burdett and Mortensen, 1998), but they may also reflect other mechanisms, such as match-specific productivity or compensating differentials (Gibbons et al., 2005; Sorkin, 2018). The underlying mechanisms matter for how we interpret the firm pay premia $\psi_j$ estimated by the AKM model, which corresponds to pairing eq. (1) with

[6] It is important to account for education-specific variation in earnings over age when studying career dynamics by parental income (Section 3), since education and parental income are positively correlated.

a strict exogeneity restriction on the errors $\varepsilon_{ijt}$. As discussed by Kline (2024), this restriction embeds several assumptions, such that worker mobility between firms is not driven by time-varying wage shocks ("exogenous mobility"), that firm pay premia are stable over time ("no drift"), and that worker and firm heterogeneity enter log earnings additively.[7]

We estimate equation (1) using our AKM 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 men.[8] As the firm pay premia $\psi_j$ in equation (1) are identified from earnings changes when workers switch firms, they are identified only relative to a baseline firm within a set of firms connected through such firm-to-firm transitions ("movers"). Firm fixed effects from disconnected sets are not comparable. We therefore follow standard procedures to compute the largest connected set for our time period and drop worker-year observations associated with firms outside this set (about 0.7% of observations).

Still, firm effects tend to be imprecisely estimated for firms connected by only a small number of movers (Andrews et al., 2008; Kline et al., 2020; Bonhomme et al., 2023), which inflates the variance of firm effects and induces a downward bias in the covariance between worker and firm effects ("limited-mobility bias"). We therefore drop firms connected to other firms by fewer than five movers, removing an additional 6% of worker-year observations.[9] This adjustment, combined with the use of population-wide data over a long time period, reduces concerns about limited-mobility bias in the variance decomposition. Moreover, because the estimated firm effects serve as the dependent variable in our intergenerational regressions, classical measurement error would not affect the consistency of our coefficients of interest.

In the subsequent sections, we use estimates of $\alpha_i$ and $\psi_j$ as inputs in our intergenerational analysis. The worker effects $\alpha_i$ represent the portable component of earnings across employers, reflecting a worker's productivity but possibly also other permanent factors, such as the worker's bargaining power or discrimination at the market level (Kline, 2024). The firm effects $\psi_j$ represent a non-portable component of earnings enjoyed only when employed at firm $j$, reflecting differences in firm wage-setting that could stem from productivity ("wage

[7]These assumptions are unlikely to be fully met in practice. Nevertheless, the linear specification in eq. (1) can serve as a useful approximation, especially given evidence that firm pay premia are relatively stable over time (Lachowska et al. 2023, Engbom et al., 2023), and that non-separable match effects between workers and firms account for only a small share of the variance in log earnings (Bonhomme et al. 2019, Kline 2024). In particular, Engzell and Wilmers (2025) show that in Sweden, the earnings boost from working at a high-paying firm is similar for workers from high and low parental earnings backgrounds.

[8]We also provide robustness tests using data on wage rates instead of annual earnings for a large subsample covering the same time period (see Appendix A1.2).

[9]While estimated firm effects are mostly used as a dependent variable, some analyses use it as a covariate (e.g. Section 5), motivating our choice to exclude firms with few movers throughout. However, Table A5 in Appendix A1.2 shows that the results remain largely similar when we include firms with fewer than 5 movers – if anything the role of firm sorting increases.

pass-through"), amenities ("compensating differentials"), or external constraints (e.g., collective bargaining agreements). Intergenerational earnings persistence may operate through either the worker- or firm-specific components of earnings.

Using our estimates from equation (1), we decompose the variance of earnings into its different components; Appendix A1.1 reports results for both the AKM and the intergenerational samples. Worker effects explain 30-38% of the variance of log earnings, firm effects explain 8-13%, and their covariance contributes another 7-12%, reflecting strong sorting of more productive workers into better-paying firms ("assortative matching"): indeed, the implied correlation coefficient between the estimated worker and firm fixed effects is between 0.40 and 0.55, depending on sample. Overall, our decomposition results are similar to Engbom et al. (2023), who use similar data and specifications, and broadly consistent with evidence from the US and other countries (e.g. Song et al., 2019; Bonhomme et al., 2023).

## 2 The contribution of firms to intergenerational persistence

We start by documenting how children from different socioeconomic backgrounds sort into firms with different pay premia, and how this sorting contributes 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 $\beta$ the intergenerational earnings elasticity (IGE) .

Parental background may influence child earnings through both the portable worker and the non-portable firm components of the AKM model. For example, parental investments in a child's human capital would primarily be reflected in the worker effects $\alpha_i$. Conversely, preferential labor-market access through parental networks may manifest as a positive association between parental earnings and the firm effects $\psi_j$. Because $y_{ijt}$ can be decomposed according to equation (1), the slope coefficients from separate regressions of each of its components on $y_{f(i)}$ sum to $\beta$. Our primary focus is on the slope coefficient from regressing the child's estimated firm pay premium $\hat{\psi}_j$ on parental income $y_{f(i)}$, which we denote $\beta_{\mathit{firm}}$ and refer to as the SES gradient in firm pay (or simply, the *firm pay gradient*).

Table 2 reports estimates of the IGE and its components, with child earnings and firm premia measured as averages over age 39-41. As shown in column (1), the IGE is roughly 0.20, consistent 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).[10] Columns (2) and (3) decompose the IGE into individual and firm components.[11] Almost 60% of the IGE is attributed to the individual effects, which capture all permanent determinants of earnings, such as time-constant abilities or skills. The firm effects account for 33% of the IGE, rising to 36% when netting out the AKM covariates and residuals from $y_{ijt}$.[12] The firm pay gradient thus explains an important share of the intergenerational persistence of income from one generation to the next.

*Comparison to previous studies.* Worker-firm sorting accounts for a much larger share of *intergenerational* than of *cross-sectional* inequality. Worker-firm sorting explains 7-12% of the variance in log earnings, depending on sample (see Appendix A1.1).[13] This contribution is modest relative to the individual component, representing less than a third of its size. By contrast, worker-firm sorting explains 36% of the IGE in our data, more than half of the individual component.

We find a larger role for firm sorting than Dobbin and Zohar (2023), who estimate that firms account for about 23% of the IGE in Israel (after netting out the AKM residuals) – roughly two thirds of the share that we find in our setting. Moreover, in their estimates the individual component is far more dominant: worker-firm sorting contributes only 29% relative to the individual component, compared with 56% in our analysis. The firm contribution we find is also larger than in Engzell and Wilmers (2025), who use similar Swedish data but focus on earnings ranks rather than levels.

Why do we find stronger sorting of workers across firms by family background in Sweden than in previous work from Israel? One possibility is that firm sorting is genuinely stronger in Sweden. For instance, Sweden has relatively low returns to skill and a homogeneous, tuition-free education system, which may raise the *relative* importance of firm sorting compared to time-invariant skills in the intergenerational transmission of earnings. But another possibility is that differences in the quality of the underlying income data explain part of the gap. Dobbin and Zohar observe Israeli workers over a six-year window (2010-2015), while our data cover the Swedish labor force for a much longer period (1985-2018), allowing for more precise estimation of firm pay premia.

[10] 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.

[11] While the coefficients in a second-step regression of the estimated firm effects on worker covariates are unbiased, inference is complicated by correlation between the fixed effects estimates (Kline, 2024). In an example application based on two years of income data, Kline (2024) shows that standard errors based on the cross-fitting procedure proposed by Kline et al. (2020) are about 75% larger than the naive standard errors. This bias is less of a concern here, given the small size of the standard errors reported in Table 2, and the larger number of person-year observations used to to estimate firm effects in our AKM sample.

[12] The correlation between the log of father's income and the fitted values from the time-varying controls (year and education-by-gender-specific age dummies) and the residuals, $\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$, explains 9% of the IGE (see column 4). Abstracting from this component, firm effects explain 33%/(59%+33%) = 36% of the IGE.

[13] This is similar to estimates from other countries. For example, Bonhomme et al. (2023) report bias-corrected estimates of the firm contribution to inequality ranging between 5-20% for the US and four other countries, including Sweden.

Table 2: Decomposition of the intergenerational earnings elasticity

| | Dependent variable | | | | |
|---|---|---|---|---|---|
| | $y_{ijt}$ | $\hat{\alpha}_i$ | $\hat{\psi}_{j=J(i,t)}$ | $\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$ | $\hat{\psi}_{j=J(i,t)}$ |
| | (1) | (2) | (3) | (4) | (5) |
| $y_{f(i)}$ | 0.198 | 0.116 | 0.065 | 0.018 | 0.047 |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| $\hat{\alpha}_i$ | | | | | 0.156 |
| | | | | | (0.001) |
| Share of IGE | 100% | 59% | 33% | 9% | 24% |
| Worker obs. | 836,743 | 836,743 | 836,743 | 836,743 | 836,743 |

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 $\alpha_i$, mean firm fixed effects $\psi_j$ for age 39-41, and time-varying controls. Robust standard errors in parentheses.

When we replicate their data scenario and restrict attention to earnings records from 2010-2015, we find an SES gradient in firm pay that drops substantially from 0.065 to 0.038 (see Appendix Table A3), and explains only 18% of the IGE (20% when netting out the AKM covariates and residuals). Classical measurement error would not affect this estimate, since firm pay is the dependent variable. However, Figure A1 shows that the error is not classical: firm premia estimated from the short panel systematically overstate firm premia at low values and understate at high values along the firm premium distribution. Differences in the quality of the AKM estimates may therefore drive part of the cross-country discrepancy in the firm contribution to the IGE.[14] Section 2.3 provides further evidence on the sensitivity of our estimates to sampling and specification choices.

*Skill Sorting.* While worker-firm sorting explains a substantial share of the IGE, it may be only indirectly related to family background. Prior work documents strong assortative matching between workers and firms, such that individual and firm fixed effects are positively correlated (Card et al., 2013; Song et al., 2019). This sorting contributes to income inequality in the cross-section, as shown in eq. (A1) and Appendix Table A1. It also contributes to intergenerational persistence, since worker effects are positively correlated with parental income (Table 2, column 2). Part of the observed SES gradient in firm pay could therefore arise as a "mechanical" consequence of this sorting of more productive workers to higher-paying firms.

[14] 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 other components in the AKM estimation. Randomly dropping half of our sample has little effect on our estimates (Appendix Table A4). Excess homogeneity from non-random sampling 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.

To illustrate the potential contribution of assortative matching, we follow Dobbin and Zohar (2023) and regress firm fixed effects on father's log income *conditional* on worker fixed effects (Table 2, column 5). The coefficient on the worker effects is large and statistically significant, reflecting substantial assortative matching, while the coefficient on father's log income declines by nearly 30%, to 0.047. Part of the firm pay gradient thus originates from the SES gradient in permanent worker characteristics, as captured by the individual fixed effects. Put differently, worker-firm sorting magnifies the earnings impact of the well-known SES gradient in human capital by amplifying the pay-off to skills. Yet, the remaining SES gradient in firm pay still explains 24% of the IGE, suggesting that sorting by family background across firms is substantially stronger than would be implied by skill sorting in the standard AKM framework. Skill sorting is difficult to measure precisely, however, and we return to this issue in Section 4.

## 2.1 Non-linear firm pay gradient

Does the sorting across firms matter more among low- or high-income families? Figure 1 plots expected child earnings (subfigure a) and firm premia (subfigure b) at age 40 across the distribution of parental income, with incomes now expressed in ranks (see Figure A2 in the Appendix for the corresponding results using log income). Both figures reproduce the positive intergenerational relationships documented in Table 2, but reveal that the gradient steepens sharply above roughly 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 gap between the unconditional (blue circles) and conditional (red triangles) gradients widens toward the top of the parental-income distribution, indicating that skill-based sorting plays a particularly important role among higher-income families. Taken together, these results suggest that sorting into high-paying firms is especially important for intergenerational earnings transmission among children from affluent families.

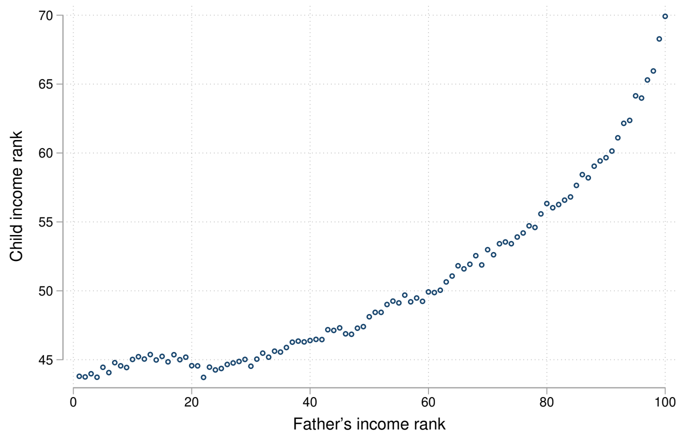
## 2.2 Region, industry, and firm size in the firm pay gradient

Pay premia vary not only across firms, but also across regions and industries. In our data, the industries with the highest average premia include oil and natural gas extraction, financial-sector support services, telecommunications manufacturing, chemical manufacturing, IT services, and research and development. Moreover, larger firms tend to pay higher premia than smaller firms. The firm pay gradient documented in Table 2 may therefore reflect the sorting of high-SES children certain industries or into particular regions, industries, or larger firms.

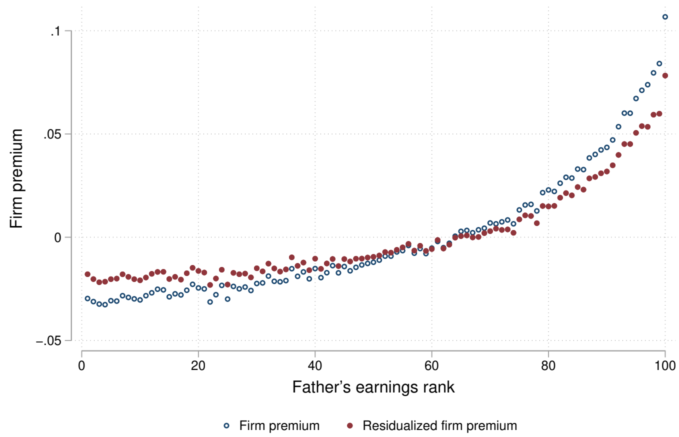
Table 3 explores this possibility by sequentially controlling for various sets of firm char-

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

(a) Child income rank at age 40




(b) Firm premia at age 40




Notes: Binned scatter plots by father's income rank. Subfigure (a) shows child's average income rank at age 40. Subfigure (b) shows average firm premia $\psi_j$ at age 40 as estimated by equation (1) and firm premia residualized on individual fixed effects.

acteristics. Controlling for region (21 counties, column 2) reduces the SES gradient in firm pay by about 30%. Restricting the analysis to variation within 2-digit industries (59 in total, column 3) yields an even larger reduction of roughly 33%.[15] Controlling for sector (public vs private) has only a minor impact (column 4), while conditioning on firm size has a moderate effect on the estimated SES gradient, reducing it by 17% (column 4). When jointly controlling for region, industry, sector, and firm size, about 39% of the unconditional SES gradient in firm pay remains unexplained. Thus, an important share of the "firm" pay gradient reflects observable firm characteristics, though a substantial share remains unaccounted for.[16]

[15] For comparison, Card et al. (2024) find that one third of the variance in firm wage effects can be explained by four-digit industry codes in the US.

[16] Of the remaining share, about half is explained by pay differences across occupations (Appendix Table A8).

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) | (6) |
| $y_{f(i)}$ | 0.064 | 0.045 | 0.043 | 0.063 | 0.053 | 0.025 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Region FEs | | X | | | | X |
| Industry FEs | | | X | | | X |
| Public sector | | | | X | | X |
| Firm size (deciles) | | | | | X | X |
| Share of $\beta_{firm}$ | 100% | 30% | 33% | 2% | 17% | 61% |
| Worker obs. | 778,795 | 778,795 | 778,795 | 778,795 | 778,795 | 778,795 |

Notes: Column (1) reports estimates of the slope coefficient from regressing $\hat{\psi}_j$, at age 40 as estimated from equation (1), on log fathers earnings. Columns (2)-(6) report coefficient estimates from the same regression controlling for region fixed effects (21 counties), industry fixed effects (2-digit level, 59 industries), working in the public sector (public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99), or log establishment size. Robust standard errors in parentheses.

## 2.3 Sensitivity and heterogeneity

We conduct several additional analyses to examine subgroup differences and the sensitivity of our main estimates to specification and sampling choices. First, to address concerns that the long estimation window may violate the assumption of time-invariant firm pay premia, we estimate a time-varying AKM model (e.g. Lachowska et al., 2020; Engbom et al., 2023) by splitting the 1985-2018 period into four subperiods. Allowing firm effects to vary over time modestly increases the estimated role of firms (Appendix Table A5, Panel A).

Second, we consider sorting across establishments rather than firms. Establishment fixed effects explain 35% of the IGE (39% after netting out AKM residuals, Table A5, Panel B), slightly more than with our baseline firm definition.[17] Using firm identifiers for all workers instead – regardless of firm size – reduces the estimated firm contribution, largely due to public-sector employment (Table A5, Panel C).

Third, we vary the minimum number of "movers" required for each firm in the AKM estimation. Restricting the sample to firms for with at least 10 or at least 50 movers reduces the firm pay gradient (Table A5, Panels E and F) compared to retaining all firms (Panel D). While excluding firms with few movers mitigates limited-mobility bias (Bonhomme et al., 2023), firm fixed effects may be less informative about worker-specific pay in large firms. We therefore retain our baseline specification, which retains firms with at least five movers.

Fourth, we consider monthly full-time wages instead of annual labor earnings, considering the subsample of workers covered by an employer survey. While the estimated firm contribution is notably smaller (Appendix Table A6), this partly reflects sample selection, as

[17] Note that our baseline already separates very large firms into establishments (see Section 2) and that many smaller firms consist of only one establishment. Overall, using establishments increases the number of "firm units" from XX to YY.

the wage sample oversamples large firms and public-sector workers. We therefore focus on earnings in our main analysis, 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 across subgroups. The firm pay gradient as a share of the IGE is very similar for women and men (Appendix Table A7, Panels A and B). And while a non-negligible share of children works in the exact same firm as their parents (in line with Corak and Piraino, 2010; Stinson and Wignall, 2018; Laliberté and Whalley, 2025), dropping these "firm followers" leaves the firm pay gradient virtually unchanged (Panel C). In contrast, excluding public sector workers increases the firm contribution by nearly 10% (Panel D), consistent with more compressed wage structures and the presence of large public employers, for which firm fixed effects may not well reflect pay in different groups within the "firm".

# 3 Firm pay gradients over the lifecycle

We found that by age 40, children from high-income families are more likely to work at higher-paying firms, that this pattern holds conditional on their own permanent skills as captured by the individual fixed effects, and that such firm-sorting accounts for a substantial share of the persistence of income inequality across generations. This section turns to the career dynamics underlying these patterns, examining when and how individuals from more privileged family backgrounds come to be employed at higher-paying firms.

## 3.1 Firm pay over the lifecycle

Figure 2 plots age profiles of the mean firm pay premium by quartile of parental income. Panel (a) shows a pronounced SES gradient already at age 25, with children from higher-income families working at firms with higher pay premia. Moreover, individuals from the top parent-income quartile (yellow squares) experience a notably steeper gain in firm premia early in their careers. While children from the bottom three quartiles see their firm premium rise by about 1 percentage point up to their early 30s, the corresponding increase among those from the top quartile is around 3 percentage points. The gap stabilizes from the mid-30s onward. Given a standard deviation of firm premia of 0.14-0.15 (see Table A1), the advantage of the top quartile relative to the next quartile corresponds to about 29% of a standard deviation. These lifecycle patterns are qualitatively very similar for men and women (see Appendix Figure A6), though average firm pay premia are consistently higher for men.

Some of these dynamics may reflect education-specific differences in career trajectories, given higher educational attainment and later labor-market entry among children from high-income families. To address this, Figures 2b and 2c reproduce the analysis separately for children with at most a high school degree and those with some college or more. In both groups, a clear SES gradient in firm pay is evident. However, among non-college workers, firm premia tend to rise less steeply over age. Among the college-educated, premia increase across all quartiles, but the gains are notably larger for those from the top parental-income quartile. Despite differences in levels and growth rates, the qualitative patterns are strikingly similar across education groups.[18]

Given this robustness, we focus on the results for the full sample. Education-specific results are, however, reported in Appendix A3 and discussed in the main text where relevant. Finally, Appendix Figure A5 plots the lifecycle evolution of firm pay premia conditional on individual fixed effects, accounting for assortative matching between workers and firms due to skill sorting (the lifecycle analogue of column 5 in Table 2). The resulting gaps in firm pay premia – and in particular the pronounced early-career increase among children with high-income fathers – remain largely similar.

In sum, gaps in firm pay emerge already at labor market entry, with high-SES children substantially more likely to work at firms that, under the standard AKM assumptions, pay otherwise identical workers more. This gap widens further during early careers before stabilizing by the mid-30s. In Section 4, we show that the initial gap cannot be accounted for by differences in worker skills, while the subsequent widening of the gap largely reflects skill differences that correlate with parental background.

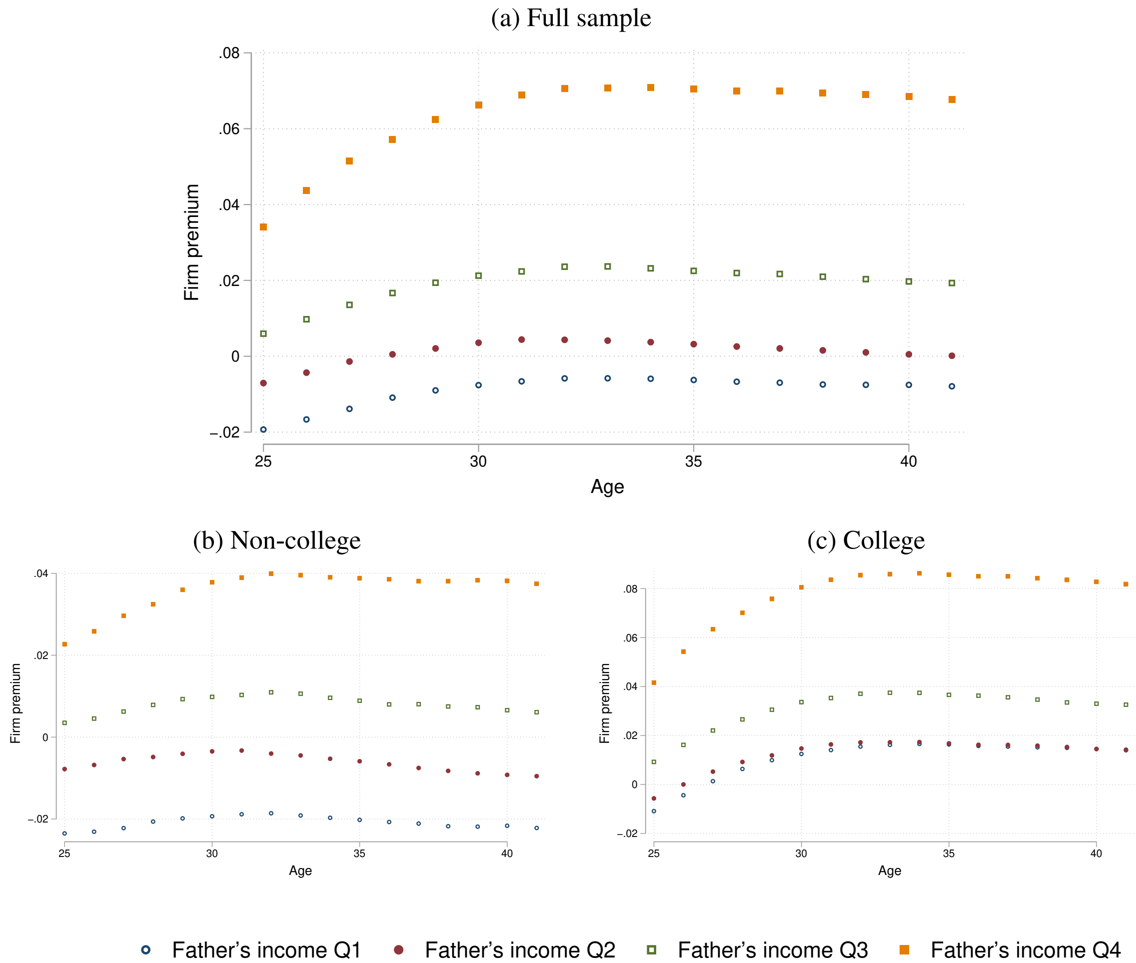
## 3.2 Climbing the job ladder

While high-SES children tend to start their career at higher-paying firms, a substantial part of the firm pay gradient develops during the first 10-15 years in the labor market. This widening may reflect either more frequent firm switches among high-SES individuals (i.e., climbing the job ladder faster), more advantageous switches (i.e., the rungs of their ladder are further apart), or a combination of both.

Figure 3a shows the annual firm-switching rate, i.e. the probability of being observed in a new firm at age $a$ relative to age $a-1$, by age and parental-income quartile. As expected, firm switching is more common in early career: roughly 22-28% of workers change firms each year between ages 25 and 30, compared to only 15% at age 40. Up to the early 30s, high-SES children switch firms more often than others, though these differences fade after age 35. This

[18]This similarity reflects that time-constant differences in worker pay are captured by the individual fixed effects, while differences in lifecycle growth *within* firms are captured by the education-age interactions in equation (1). To further probe that our results on career dynamics are not driven by differences in age at labor-market entry, we reproduce the analyses by potential experience, finding very similar results (see Appendix A3).

Figure 2: Firm earnings premium over the lifecycle

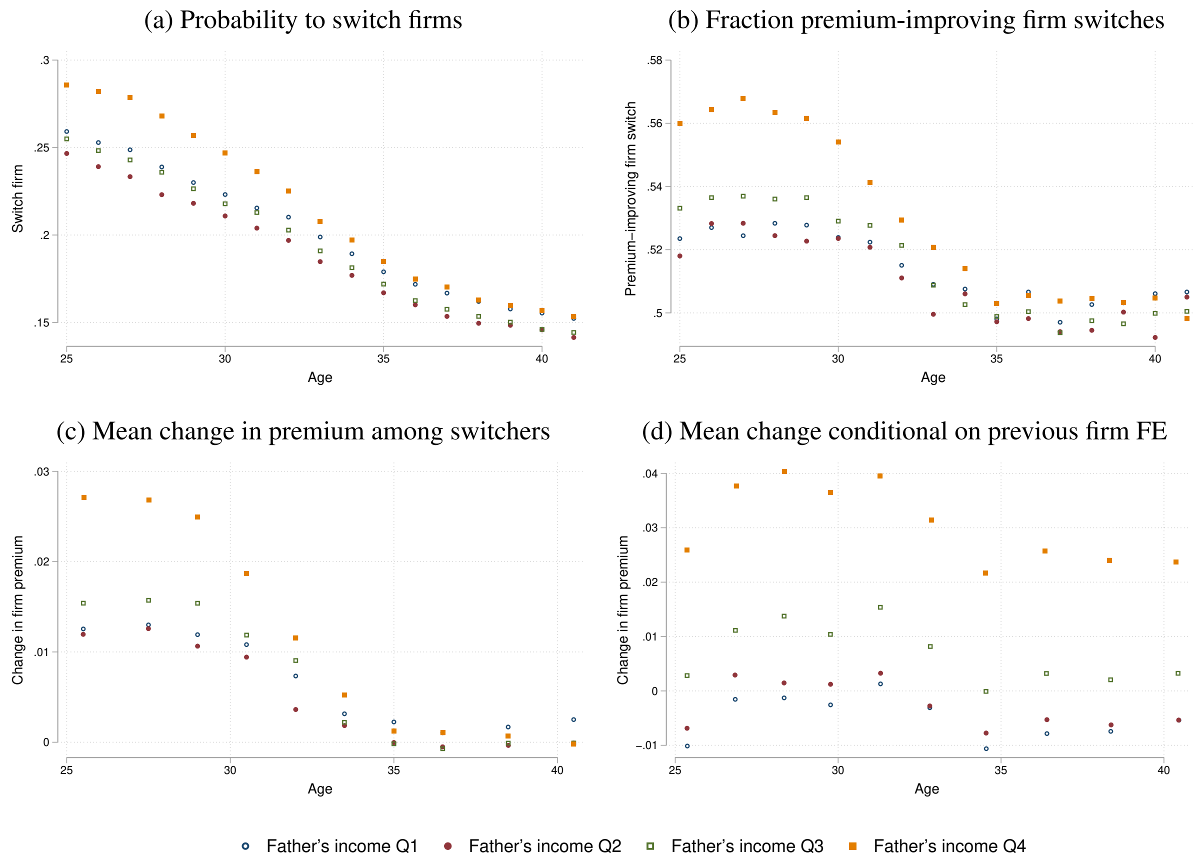



Notes: Mean estimated firm premium $\hat{\psi}_j$ over the life cycle, by quartile of father's income. Subfigure (a) shows the results for the full sample, subfigure (b) for children without college education, and subfigure (c) for children with college education. Fathers' income quartiles are defined in the full sample.

pattern holds within education groups (Appendix Figure A10), with higher mobility overall among the college educated and a more pronounced SES gradient within that group.

Switching firms, however, does not necessarily imply moving to a better-paying employer. Figure 3b therefore plots the share of firm switches that are "premium-improving" – defined as moves to a firm with higher estimated pay premium than the previous employer. All groups tend to switch to better-paying firms early in their careers (i.e., they "climb the firm ladder"), but high-SES children are markedly more likely to make such favorable moves. At later ages, the likelihood of good and bad switches tend to even out, and the SES gradient diminishes. Notably, while low-SES children switch firms slightly more often than middle-SES peers (subfigure a), they are not more likely to gain in firm pay conditional on switching. This pattern suggests that high- and low-SES children switch for different reasons, perhaps because the latter experience more involuntary switches or move between temporary jobs.

Figure 3: Firm switching patterns over the lifecyle




Notes: Patterns of firm switching over the lifecyle, by quartile of father's income. Subfigure (a) shows the probability of working in a different firm than in the year before. Subfigure (b) shows, conditional on switching, the probability of moving to a firm with a higher firm premium. Subfigure (c) shows, for individuals who switch firms, the difference between the firm premium at the new and previous firm. Subfigure (d) shows the same but adds a control for the firm premium in the previous firm.

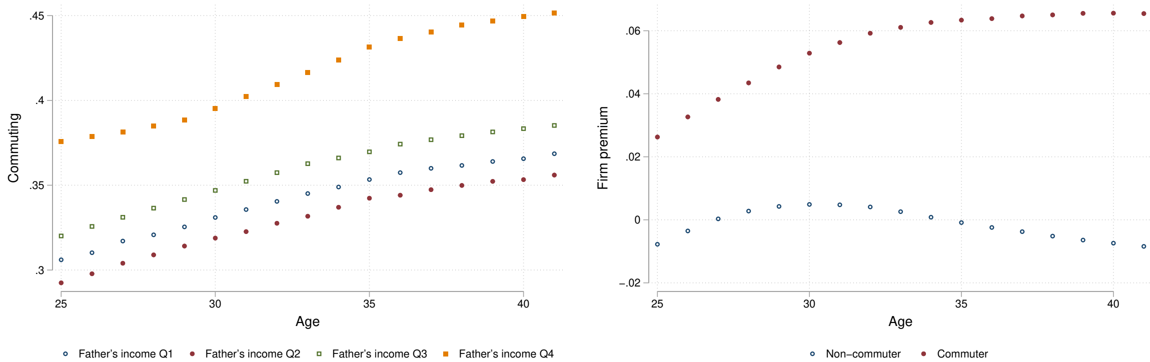
Figure 3c shows that on average, high-SES children raise their firm pay by nearly three percentage points with each switch, more than twice as much as those from lower quartiles.[19] This difference narrows with age, but this may reflect a mechanical effect: individuals who have already climbed to high-paying firms have less scope for further improvement than those starting lower ("ceiling effects"). In a random-switching scenario, low-SES workers with lower initial firm premia would experience larger average gains. To account for this, Figure 3d reports changes in firm premia conditional on the firm effect in the previous year. With this adjustment, high-SES children continue to experience systematically larger gains from firm switching throughout the life cycle, confirming the importance of ceiling effects.

Appendix Figures A11 and A12 replicate these analyses separately by education group. College-educated individuals are generally more likely to experience premium-improving

[19] The larger gains from firm switching among high-SES children cannot be explained by unemployment dynamics. To show this, we 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 A13, high-SES children are more likely to experience improvement in the firm premium following both voluntary and involuntary switches.

Figure 4: Commuting patterns

(a) Commuting rates by parental income (b) Mean firm premia by commuting status




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

switches and larger gains in firm pay throughout their careers, yet the SES gradients are qualitatively similar across education groups.

## 3.3 Commuting up the job ladder?

One potential explanation for the more frequent job among high-SES children is greater job mobility. In particular, broader social networks and stronger support structures – such as access to childcare, flexible transportation, or remote work options – may reduce the costs of commuting, thereby expanding the set of feasible jobs. By widening job opportunities, commuting can raise wages and firm pay premia (Le Barbanchon et al., 2020; Agrawal et al., 2024). At the same time, commuting is costly, so firm pay differences associated with commuting may overstate welfare differences.

Consistent with this mechanism, Figure 4a shows that children from high-income families are substantially more likely to commute (defined as working in a different municipality than their residence). At age 25, the commuting rate is 8 percentage points higher in the top than in the bottom parental-income quartile, rising to 10 points by age 40. These patterns persist within education groups (Appendix Figure A15), indicating that education alone does not explain the SES gradient in commuting.

Figure 4b shows that commuters work at higher-paying firms, especially later in the career: at age 40, the firm pay gap between commuters and non-commuters is about 6 percentage points, roughly half a standard deviation. Much of this difference, however, reflects selection. Event-study estimates conditioning on individual fixed effects (Appendix Table A11) show that starting to commute increases firm pay premia by only about 1-2 percentage points. Thus, although high-SES individuals commute more, commuting explains only a modest share of the SES gradient in firm pay.

## 3.4 Beyond firm pay premia

Firms differ along many important dimensions beyond pay. In this subsection, we study whether the firms in which children from high-income families work provide other advantages aside from higher pay premia. We begin with "static" firm characteristics related to workforce composition. Figure 5a shows that high-SES children tend to have more productive co-workers, with mean AKM effects about 6 percentage points higher for children in the top than in the bottom parental-income quartile.[20] This segregation of workers by productivity is relatively stable over age and, if anything, more pronounced in early career. If more productive co-workers increase one's own productivity (Kremer, 1993), such segregation may also contribute to the firm pay advantages documented in Section 3.1.

Sorting into firms with more productive workforces may shape long-run career trajectories if, for example, these environments generate positive learning spillovers or help build valuable social networks. Moreover, firms may offer different opportunities for career development, partly because firms themselves grow at different rates.

To illustrate that firms differ in such dynamic sense, subfigures b-d in Figure 5 show mean co-worker employment and earnings growth over the subsequent five years (from age $a$ to age $a+5$). Figure 5b shows that high-SES children are more likely to work in firms with faster employment growth: children in the top parental-income quartile work in firms that grow 40-100% faster than those employing children in the lower quartiles. They also work in firms with faster earnings growth. Figure 5c plots the mean five-year earnings growth for co-workers who stay in the same firm ("stayers"), while Figure 5d shows the corresponding growth for all co-workers, regardless of whether they stay or leave the firm. The SES gradient is much more pronounced for the latter, suggesting that much of the difference in co-worker earnings growth arises from mobility to other firms, rather than differential earnings growth among stayers. This finding aligns with earlier results that high-SES children cluster together, switch firms more frequently, and make more advantageous job moves.[21]

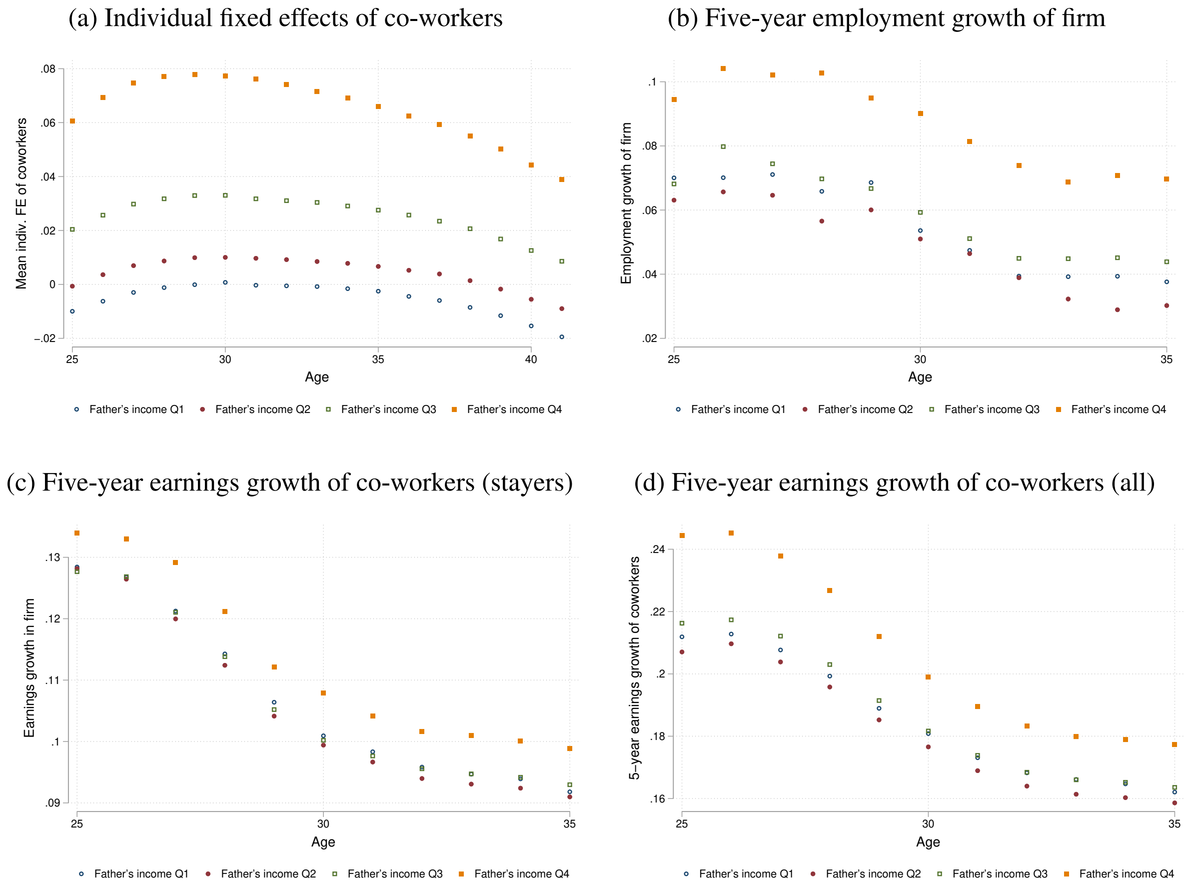
## 3.5 Heterogeneity in firm-specific returns to experience

The preceding results showed that children from high-income families are more likely to sort into growing firms and to work alongside more productive co-workers whose earnings rise more rapidly. These findings suggest that such firms may not only provide "static" pay advantages – as captured by the AKM framework – but also better opportunities for career development in a "dynamic" sense. For example, faster co-worker earnings growth may

[20] Unsurprisingly, they are also exposed to a higher share of high-SES co-workers (Appendix Figure A9).

[21] Interestingly, gaps between the top parental income quartile and the rest in co-worker earnings growth and firm employment are pronounced and relatively stable over age, while differences between those from the first three quartiles are generally much smaller.

Figure 5: Static and dynamic firm characteristics




Notes: Subfigure (a) shows the mean individual fixed effect of co-workers, by quartile of father's income. The other subfigures show the mean difference in firm characteristics between year $t$ and year $t+5$ for the individuals who work in the firm at each indicated age. Subfigure (b) plots the five-year employment growth in the firm. Subfigure (c) plots the five-year earnings growth for coworkers who stay in the firm, while subfigure (d) plots the earnings growth for all coworkers, including those who stay and those who switch firm.

reflect better learning or promotion opportunities. Even firms that do not pay particularly well at a given point in time could be attractive if they offer high returns to accumulated experience, thereby increasing earnings later in the career.

To examine the role of such firm-specific returns to experience, we follow Arellano-Bover and Saltiel (2021). We begin by randomly splitting our sample into two groups. Using one of the subsamples, we classify firms into ten classes based on the distribution of stayers' annual unexplained earnings growth, using a $k$-means clustering algorithm. We then use the second subsample to estimate firm-class-specific returns to experience. In particular, we estimate an extended two-way fixed effects model:

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

where $Exp(m)_{it}$ denotes years of experience in firm class $m$ up until year $t$. As above, we include individual and firm fixed effects, such that $\gamma_m$ is identified from workers employed in the *same* firm but with different experience histories across firm classes. As before, $\boldsymbol{X}_{it}\boldsymbol{\beta}$ includes age-education-gender controls and year fixed effects.[22]

Figure 6 summarizes the results. Panel (a) shows the estimated returns to an additional year of experience in each firm class $m$, where class 1 contains firms with the highest returns and class 10 those with the lowest. Similar to Arellano-Bover and Saltiel (2021) and Battiston et al. (2024), we find substantial heterogeneity across firm types. However, the differences in firm-specific returns in our Swedish data is smaller than in Italy or (especially) Brazil: workers in the top firms (class 1) experience an annual earnings growth of about 2.5 percentage points, while those in class 10 experience slightly negative returns (net of general education-by-gender specific earnings growth, as captured by $\boldsymbol{X}_{it}\boldsymbol{\beta}$).

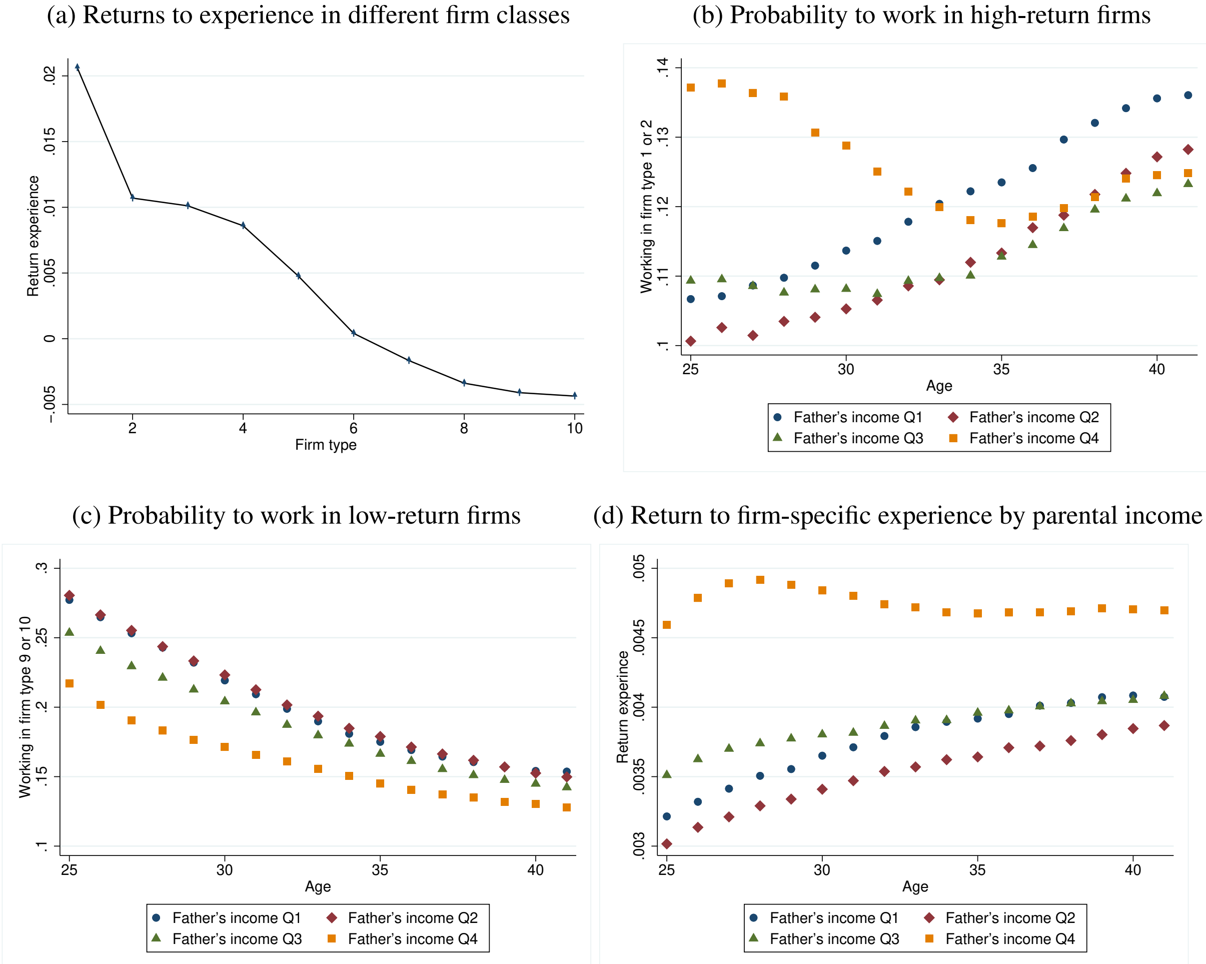
Figures 6b and 6c plot the probability to work in high-return firms (classes 1 or 2) and low-return (classes 9 or 10) firms by parental-income quartile and age. Children from high-income families are more likely to work in firms with very high returns to experience early in the career and less likely to work in low-return firms. The SES gap is substantial at younger ages, when about 14% of top-quartile children and about 10-11% of the other children work in high-return firms, but shrinks with age. Moreover, top-quartile children are less likely to work in low-return firms at all ages. Differences among the bottom three quartiles are smaller; if anything, children from the bottom quartile are somewhat more likely to be employed in both high- and low-return firms.

Figure 6d shows the estimated average firm-specific return at an individual's current employer by age and parental-income quartile. Children from the top quartile consistently work in firms with higher returns to experience, with the SES gap widening up to around age 28 and narrowing slightly thereafter. In magnitude, those from the top parental-income quartile enjoy firm-specific returns that are on average about 40% higher up to age 30 compared to those of their lower-SES peers. Interestingly, it is those from the second-to-lowest rather than from the lowest quartile who experience the lowest firm-specific returns.

Table 4 quantifies how much of the intergenerational earnings elasticity at age 40 can be attributed to the components in equation (3). The decomposition separates the firm pay premium at age 40 into the early-career firm premium at age 25 (column 3) and the *change* in firm pay between ages 25 and 40 (column 4). About 40% of the SES gradient in firm pay at age 40 – and roughly 16% of the IGE – can be attributed to changes in firm pay over age (i.e., climbing the firm ladder, see Section 3.2). Additionally, sorting into firms with higher

[22] 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. Instead we normalize age relative to age 40 and include second and third order polynomials of age interacted with education and gender (rather than age dummies).

Figure 6: 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: Subfigure (a) shows estimates of the coefficients on firm-class experience from equation (3). Subfigures (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. Subfigure (d) shows the average firm-specific returns to experience, by quartile of father's income.

*returns to experience* explains another 8% of the IGE (column 5).

Because identification of the return to firm-specific experience relies on workers currently at the same firm but with different prior experience, 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 A16b in Appendix A4 shows a positive relationship between estimated returns to experience $\gamma_m$ (as captured by column 5) and changes in firm premia (column 4), suggesting that firm-specific experience also helps explain why high-SES children climb the job ladder more quickly. Taken together, this dynamic perspective on firm pay premia implies an even larger role for firms in intergenerational earnings persistence than suggested by the static decomposition alone (in Table 2). Combining columns (3)-(5) suggests that firms account for about 46% of the IGE (or 50% net of column 6).

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) |
| $y_{f(i)}$ | 0.180*** | 0.080*** | 0.040*** | 0.028*** | 0.015*** | 0.017*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Share of IGE | 100% | 44% | 22% | 16% | 8% | 9% |
| Worker obs. | 253,072 | 253,072 | 253,072 | 253,072 | 253,072 | 253,072 |

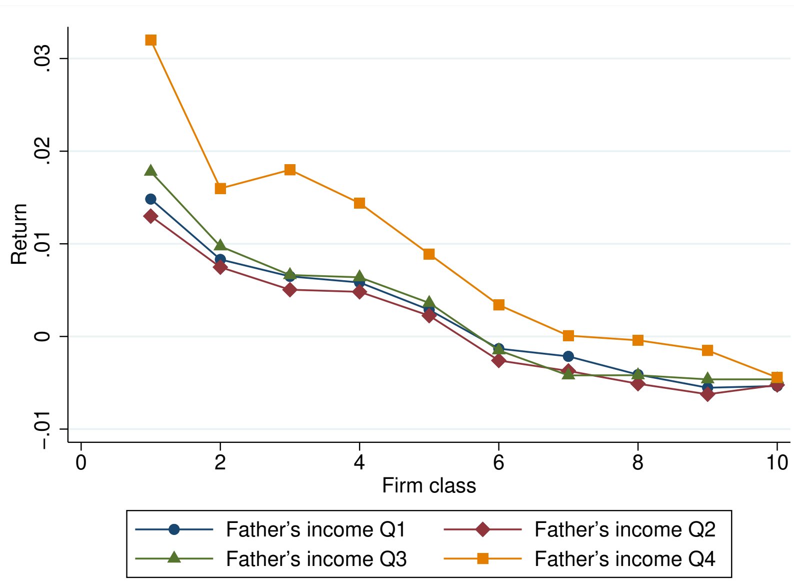
Notes: Column (1) reports the estimated slope coefficient from a regression of log child earnings at ages 40 on log father's earnings. Columns (2)-(6) report slope coefficients from the corresponding regressions when decomposing child log earnings $y_{ijt}$ according to equation (3) into the individual fixed effect $\alpha_i$, the firm fixed effect at age 25, the change in the firm fixed effect between age 25 and 40, returns 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, half of the sample is used to cluster firms into firm classes, and the sample is restricted to individuals observed in a firm at age 25. Robust standard errors in parentheses.

A potential concern is that children from high-income families may systematically experience higher returns to experience regardless of which firm they sort into. In that case, the differences in returns across firm classes as shown in Figure 6a would reflect heterogeneity across individuals rather than firms. Arellano-Bover and Saltiel (2021) address this concern by interacting firm-specific experience with worker-fixed effects. We follow a similar approach by allowing the returns to experience to vary with parental income, estimating

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

where $\theta_i$ indicates the father's income quartile. Figure 7 shows that children from high-income families have higher returns to experience within all firm classes. Nonetheless, all workers benefit from working in high-return firms, and the variation in returns across firm classes remains nearly as large as in the baseline specification – and is even larger for the top quartile (cf. Figure 6a).

Figure 7: Returns to experience in different firm classes by father's income




Notes: Estimates of the coefficients on firm-class experience from equation (4), by quartile of father's income.

# 4 The role of skill sorting in the firm pay gradient

Part of the SES gradient in firm pay reflects assortative matching between workers and firms. Children from high-income families tend to be more educated and to have higher permanent productivity – as captured by the AKM individual fixed effects – and workers with higher permanent productivity (or skills) sort into firms that pay higher premia (e.g., Card et al., 2013). Yet, roughly two thirds of the SES gradient in firm pay remains even after conditioning on the estimated individual fixed effects (see Table 2, column 4).[23] For brevity, we refer to this remaining component as the *conditional firm pay gradient*.

There are two broad interpretations for why a substantial SES gradient remains even after accounting for this proxy for individual skills. First and foremost, parental income may exert a direct effect on firm sorting above and beyond what is driven by skill differences among the children. Such direct effects could stem from several sources, including informational advantages, social and co-worker networks, and credit or other constraints.[24] They may also arise if preferences for non-monetary job attributes vary by family background, implying that part of the firm pay gradient reflects compensating differentials or other non-pay amenities offered by firms (see Section 5).

[23]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.

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

However, an alternative interpretation is that the estimated worker effects $\hat{\alpha}_i$ do not fully capture the extent of skill sorting, since they are only imperfect proxies of skill. First, the $\hat{\alpha}_i$ are noisy estimates of $\alpha_i$ in the AKM model, which biases downward the estimated contribution of assortative matching to the firm pay gradient (Dobbin and Zohar, 2023). Second, the individual effects $\alpha_i$ capture all *persistent* within-firm differences in earnings, not solely those arising from differences in productivity. For example, persistent taste-based discrimination across ethnic groups would load onto $\alpha_i$, even if underlying productivity is similar.[25] Third, skills are multidimensional, and worker-firm sorting may occur along specific dimensions (e.g., cognitive skills) rather than the entire bundle of skills embodied in the AKM individual fixed effects. If parental background correlates more strongly with some skills than with others, conditioning on the aggregate $\alpha_i$ may obscure relevant channels of sorting.

## 4.1 Measuring the contribution of skill sorting

To more directly analyze skill sorting, we incorporate explicit measures of cognitive and social (or "non-cognitive") skills obtained from military enlistment tests, which we use alongside our fixed effects-based skill proxy.[26] Specifically, we apply a decomposition similar to Gelbach (2016) and Hjorth-Trolle and Landersø (2023) to examine sorting on (i) cognitive skills, (ii) social skills, (iii) education and (iv) the estimated AKM individual fixed effect. Because the enlistment testing was mandatory only for men, we restrict this analysis to males.[27] These skill measures 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 framework allows us to test whether the AKM fixed effects–often used as controls in worker-firm sorting analyses–are incomplete measures of skill, and whether sorting is therefore underestimated. It also enables a comparison of the relevant importance of different skill dimensions (cognitive, social, etc) for explaining firm

[25] Moreover, Dobbin and Zohar (2023) note that the AKM worker effects $\alpha_i$ 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. Conversely, Bello and Morchio (2022) predict that "occupational followers" who choose their father's occupation earn lower wages, due to skill mismatch. 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.

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

[27] The military tests are taken at around age 18 and were compulsory for all men in the cohorts we study. The cognitive skill score aggregates four subtests measuring verbal, logical, spatial and technical skills. The social/non-cognitive 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 details, see e.g. Lindqvist and Vestman, 2011). Both scores are standardized to mean zero and standard deviation one within each birth year.

sorting patterns.

The decomposition, summarized by regression equations (5a)-(7), partitions the relationship between children's firm premia and parental log income into various factors influencing children's earnings. After estimating the firm pay gradient $\beta_{\mathit{firm}}$ in equation (5a), we augment the regression with our four mediators of interest: cognitive skill, social skill, education, and the AKM individual fixed effect:

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

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

The coefficient $\beta_{\mathit{firm,res}}$ in equation (5b) captures the *direct* effect of family background on firm premia not mediated by observed skills, while the difference $\beta_{\mathit{firm}} - \beta_{\mathit{firm,res}}$ captures the part explained by the mediators. To quantify how parental income relates to each skill measure, we estimate the auxiliary regressions:

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

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

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

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

The part of the firm pay gradient attributable to skill sorting is then given by

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

Table 5 presents the results, focusing on firm premia measured at age 40. Although the sample is restricted to men with observed skill measures, the baseline estimate of the relationship between firm pay and log parental income in column (1), $\hat{\beta}_{\mathit{firm}} = 0.069$, closely matches those obtained for all males (0.070, Table A7) or the full population (0.065, Table 2). Column (2) shows that sorting on our four skill measures explains nearly half of this SES gradient. Conversely, 53.6% of the gradient remains unexplained, suggesting that parental background has a substantial direct effect on firm sorting beyond what is driven by differences in children's skills.[28]

Columns (3)-(6) report the contribution of each of the skill measures to overall sorting, as captured by the $\beta\phi$ terms from equation (7). The estimated individual fixed effects and cognitive skills are both key in explaining why children from richer families work at better-paying

[28] Figure A3 shows that the firm pay gradient increases markedly along the parental-income distribution (as also visible from Figure 2b), reaching about 0.12 at high levels of parental income. The unexplained part not attributable to skill sorting also grows in absolute magnitude along the distribution, though it declines as a share of $\beta_{\mathit{firm}}$.

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

| | Firm pay premia | | | | | |
|---|---|---|---|---|---|---|
| | Overall $\beta_{firm}$ (1) | Unexp. $\beta_{firm,res}$ (2) | Cognitive $\beta_{cog}\phi_{cog}$ (3) | Social $\beta_{soc}\phi_{soc}$ (4) | Education $\beta_{edu}\phi_{edu}$ (5) | Indiv. FE $\beta_{iFE}\phi_{iFE}$ (6) |
| | 0.069*** | 0.037*** | 0.011*** | 0.001*** | 0.007*** | 0.012*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) |
| Share of IGE | 100% | 53.6% | 15.9% | 1.5% | 10.1% | 17.4% |
| Worker obs. | 361,165 | 361,165 | 361,165 | 361,165 | 361,165 | 361,165 |

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

firms, with each contributing about 16-17% to the firm pay gradient. Education is somewhat less important, explaining roughly 10%, while the role of the social skill component is substantially smaller. The limited contribution of education to the SES gradient in firm pay is surprising, given that education correlates more strongly with parental income than the other mediators (see Appendix Table A9).[29]

Table 5 thus confirms that skill sorting is underestimated – and the direct role of family background overstated – when skills are proxied solely by the estimated AKM fixed effects, as already conjectured by Dobbin and Zohar (2023). When conditioning only on worker fixed effects, 72.3% of the SES gradient remained unexplained (0.047/0.065, Table 2), or 71.4% when considering only men (0.050/0.070, Appendix Table A7). Incorporating our full set of skill measures reduces the unexplained share substantially, to 53.6% (0.037/0.069). One interpretation is that parental income influences firm sorting through mechanisms unrelated to child skills. Alternatively, employers may value certain skills that are difficult to observe but correlated with parental income.[30]

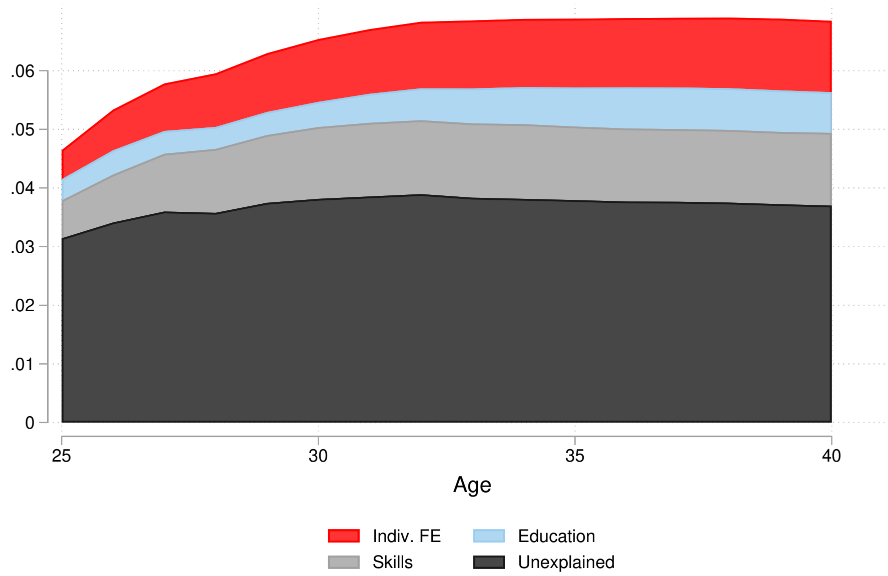
## 4.2 The role of skill sorting over the lifecycle

Does the contribution of skills to the SES gradient in firm pay vary over the child's lifecycle? Figure 2 shows that the SES gradient in firm pay emerges already at labor market entry, but continues to widen during the early career. To decompose this overall firm pay gradient into components explained by skill sorting and a residual component, we estimate equations

[29] 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.

[30] Piopiunik et al. (2020) show in a lab-in-the-field experiment that HR managers value a wide range of signals, including those indicating specific skills (such as IT or language proficiency) or social skills (such as participation in team sports). While our skill measures are unusually rich, including both cognitive and social/non-cognitive skills, they remain an imperfect proxy for the set of skills observable to employers.

Figure 8: The contribution of skills to the SES gradient in firm pay by age




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

(5a)-(7) separately for each age.

Figure 8 shows the results, revealing two key patterns. First, skill sorting accounts for only a small share of the large firm pay gradient that opens already at labor market entry. At age 25, only 30% of the SES gradient reflects skill sorting, while nearly 70% reflects direct family effects not mediated by observed skills. Second, the contribution of skill sorting, $\beta_{firm} - \beta_{firm,res}$, increases substantially over the lifecycle, while the direct effect of family background net of skill, $\beta_{firm,res}$, remains nearly constant. As a consequence, the relative importance of the two components shifts substantially: at age 25, nearly 70% of the firm pay gradient reflects direct family effects, compared to only about 50% at age 40.

The lifecycle patterns in Figure 8 seem intuitive and consistent with recent work on the importance of networks. Early in their careers, young adults remain more closely tied to their parents and parental networks, and may rely more heavily on family-provided information, referrals, or support. As workers age, their own skills become more important for firm sorting, contributing further to the overall influence of family background on the types of firms in which individuals work.

However, the sizable direct family effects net of skill sorting, $\beta_{firm,res}$, may reflect not just network effects but also statistical discrimination, if employers use family background as a proxy for unobserved skills. To illustrate this mechanism, we introduce a simple model of skill sorting inspired by the employer learning literature (Farber and Gibbons, 1996; Altonji and Pierret, 2001), which formalizes the influence of family background under imperfect information about worker skill.

## 4.3 A model of firm pay gradients under imperfect information

Assume that productivity $y_i$ depends on skills,

$$y_i = \mathit{skill}_i^* + \xi_i, \tag{8}$$

where $\mathit{skill}_i^*$ denotes an individual's true skill and $\xi_i$ is a productivity shock. Skill is increasing in socioeconomic background,

$$\mathit{skill}_i^* = \gamma_{\mathit{ses}}\, \mathit{ses}_i^* + u_i, \tag{9}$$

where $\mathit{ses}_i^*$ denotes true socioeconomic background and $u_i$ captures idiosyncratic skill components orthogonal to family background, with variances $\sigma^2_{\mathit{ses}}$ and $\sigma^2_u$, and $\gamma_{\mathit{ses}} > 0$ (as confirmed empirically in Appendix Table A9). Neither skill nor socioeconomic background is directly observed. Instead, employers observe noisy proxies,

$$\mathit{skill}_i = \mathit{skill}_i^* + v_i, \tag{10}$$

$$\mathit{ses}_i = \mathit{ses}_i^* + w_i, \tag{11}$$

where $v_i$ and $w_i$ represent classical measurement error with variances $\sigma^2_v$ and $\sigma^2_w$, respectively.

Because productivity is not directly observed, firms infer productivity from observed skill and socioeconomic background, forming the conditional expectation $E[y_i|\mathit{skill}_i, \mathit{ses}_i^*]$. Assuming linearity – either because firms use a linear approximation or because $u_i$ and $v_i$ are normally distributed so that the conditional expectation is in fact linear – this implies the population regression

$$y_i = E[y_i|\mathit{skill}_i, \mathit{ses}_i] + \varepsilon_i = b_{\mathit{skill}}\, \mathit{skill}_i + b_{\mathit{ses}}\, \mathit{ses}_i + \varepsilon_i \tag{12}$$

where the coefficients $b_{\mathit{skill}}$ and $b_{\mathit{ses}}$ can be derived using the omitted-variable formula. Finally, we assume that better-paying firms hire workers with higher expected productivity (as shown in Appendix Table A10), such that

$$\psi_{j(i,t)} = \delta E[y_i|\mathit{skill}_i, \mathit{ses}_i] = \delta\left(b_{\mathit{skill}}\, \mathit{skill}_i + b_{\mathit{ses}}\, \mathit{ses}_i\right) \tag{13}$$

This structure allows us to derive model counterparts to the total and direct firm pay gradients defined in equations (5a) and (5b). Regressing firm pay premia $\psi_{j(i,t)}$ on true socioeconomic background $\mathit{ses}_i^*$ yields

$$\beta_{\mathit{firm}} = \delta\left(b_{\mathit{skill}} \frac{Cov(\mathit{skill}_i, \mathit{ses}_i^*)}{Var(\mathit{ses}_i^*)} + b_{\mathit{ses}} \frac{Cov(\mathit{ses}_i, \mathit{ses}_i^*)}{Var(\mathit{ses}_i^*)}\right) = \delta\left(b_{\mathit{skill}}\, \gamma_{\mathit{ses}} + b_{\mathit{ses}}\right) \tag{14}$$

while regressing firm pay premia on $ses_i^*$ while controlling for $skill_i^*$ yields

$$\beta_{firm,res} = \delta b_{ses}. \tag{15}$$

*Simplified case (perfectly observed socioeconomic background)*. Consider first a simplified case, in which firms observe *true* socioeconomic background, corresponding to $\sigma_w^2 = 0$ in eq. (11). In this setting,

$$b_{skill} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \tag{16}$$

$$b_{ses} = \frac{\gamma_{ses}\,\sigma_v^2}{\sigma_u^2 + \sigma_v^2} \tag{17}$$

If skill were perfectly observed ($\sigma_v^2 = 0$), then $b_{skill} = 1$ and $b_{ses} = 0$, and family background would have no direct effect on firm pay conditional on skill. With measurement error in skill ($\sigma_v^2 > 0$), however, firms use socioeconomic background as an additional proxy for skill and $b_{ses} > 0$. More generally, the coefficients $b_{skill}$ and $b_{ses}$ decrease in own measurement error and increase in the measurement error contained in the other proxy (e.g., $b_{ses}$ decreases in $\sigma_u^2$ and increases in $\sigma_v^2$). This model can thus rationalize two key patterns from Figure 8: why the direct effect of family background (net of skills) $\beta_{firm,res}$ is positive, and why the contribution of skills to the overall pay gradient $\beta_{firm}$ increases over the lifecycle, as firms increasingly learn workers' true skills.

*General case (imperfectly observed socioeconomic background).* While rationalizing some key patterns, the simplified case with $\sigma_w^2 = 0$ cannot explain why the overall firm pay gradient $\beta_{firm}$ increases over age: plugging equations (16)-(17) into (14) yields $\beta_{firm} = \delta\gamma_{ses}$, independent of measurement error in skill. To explain the increase in the firm pay gradient, consider the general case, in which firms observe only a noisy proxy for socioeconomic background ($\sigma_w^2 > 0$). In this case, the coefficients in regression (12) are

$$b_{skill} = \frac{\gamma_{ses}^2\sigma_{ses}^2\sigma_w^2 + \sigma_u^2(\sigma_{ses}^2 + \sigma_w^2)}{\gamma_{ses}^2\sigma_{ses}^2\sigma_w^2 + (\sigma_u^2 + \sigma_v^2)(\sigma_{ses}^2 + \sigma_w^2)} \tag{18}$$

$$b_{ses} = \frac{\gamma_{ses}\,\sigma_{ses}^2\,\sigma_v^2}{\gamma_{ses}^2\sigma_{ses}^2\sigma_w^2 + (\sigma_u^2 + \sigma_v^2)(\sigma_{ses}^2 + \sigma_w^2)} \tag{19}$$

and the firm pay gradient $\beta_{firm}$ increases as skill becomes more precisely observed. In the extreme case, if skills are unobserved ($\sigma_v^2 \to \infty$), we have $\lim_{\sigma_v^2\to\infty} b_{skill} = 0$ and $\lim_{\sigma_v^2\to\infty} b_{ses} = \gamma_{ses}\frac{\sigma_{ses}^2}{\sigma_{ses}^2+\sigma_w^2}$, so that from equation (14) we have $\beta_{firm} = \delta\gamma_{ses}\frac{\sigma_{ses}^2}{\sigma_{ses}^2+\sigma_w^2}$. Conversely, when skill is perfectly observed ($\sigma_v^2 = 0$), $b_{skill} = 1$ and $b_{ses} = 0$ , and $\beta_{firm} = \delta\gamma_{ses}$. The intuition is that when firms observe family background imperfectly, they underestimate the average productivity of high-SES workers. As workers age and their skills are revealed, firms update their beliefs, increasing the overall SES gradient in firm pay.

In sum, the model explains central empirical findings: (i) both skills and family background net of skills contribute to the firm pay gradient (i.e., $\beta_{firm} - \beta_{firm,res} > 0$ and $\beta_{firm,res} > 0$), (ii) the direct effect of family background opens up already at career start, (iii) the contribution of skills to the firm pay gradient increases over the lifecycle ($\beta_{firm,res}$ increases), and (iv) the overall SES gradient in firm pay rises over the lifecycle ($\beta_{firm}$ increases). These implications follow directly from the assumption that firms observe only proxies for family background and worker productivity.

Our point here is not that networks and social connections do not matter; indeed, a growing literature documents their importance.[31] Instead, our point is twofold. First, direct parental background effects, $\beta_{firm,res}$, may reflect the influence of both networks and imperfect information, helping to understand their large magnitude. Second, even in the absence of network effects, family background can predict firm sorting and generate life-cycle patterns similar to those observed in our data. Future research could aim to disentangle the relative contributions of network effects and imperfect information to socioeconomic gradients in firm pay. While a growing literature documents the importance of networks, we are not aware of any attempts to identify the contribution of statistical discrimination to firm pay gradients.

# 5 Do firm pay premia reflect overall firm desirability?

Pay is not the only firm attribute that matters to workers. Other features – such as location, schedule flexibility, health risks and fringe benefits – also vary across employers. If such amenities are priced in the market, variation in firm pay may partly reflect compensating differentials. Sorkin (2018), leveraging job-to-job flows in US data, estimates that up to two-thirds of the variation in firm pay premia could stem from compensating differentials. Lamadon et al. (2022) similarly conclude that non-pay attributes are an important determinant of firm wage effects. Yet their work, along with related studies (e.g. Sockin, 2022; Crane et al., 2023), also finds that high-wage firms tend to offer *better* amenities, thereby widening disparities in overall compensation.

Even if pay differences understate gaps in overall compensation (pay plus amenities), however, they need not understate compensation gaps between distinct groups – in particular when those groups differ in their valuations of pay and amenities. Morchio and Moser (2024), for example, show that a substantial share of the gender pay gap in Brazil arises from gender-specific sorting on non-pay attributes, implying that pay gaps overstate welfare gaps between men and women.

[31] As discussed in more details in the introduction, contributions include Corak and Piraino (2010), Kramarz and Skans (2014), Zimmerman (2019), San (2022), Staiger (2022), Eliason et al. (2023) and Campa (2025).

We therefore study to what extent *SES gradients* in firm pay reflect differences in the overall desirability of firms, accounting for both pay and non-pay attributes. A priori, the relationship between gradients in pay and welfare is ambiguous. Children from high-income families may feel pressure to match parental earnings and thus select into higher-paying but less pleasant jobs – causing income-based mobility measures to overstate mobility in welfare. Alternatively, affluent backgrounds may provide greater flexibility to choose lower-paid but more enjoyable jobs. Indeed, Schüle (2026) finds that children from high-SES families place less value on income and job security and more value on having an interesting job; and Haeck and Laliberté (2025) show that high-SES children are overrepresented in occupations such as artists and academia, which may offer non-pecuniary benefits.

To address these issues, we study revealed-preference measures of overall firm desirability – or "firm value", denoted $V_j$ – which capture both pay and amenities. Our analysis addresses three questions: (i) whether high-paying firms are also more desirable firms, (ii) whether is there an SES gradient in firm values, and (iii) the extent to which this gradient reflects skill sorting.

Differences in firm pay *conditional* on overall firm desirability inform about SES gradients in amenities – that is, compensating differentials, or what Sorkin (2018) refers to as the "Rosen motive" for amenities.[32] However, firms may also differ in the overall utility they offer, inclusive of both pay components (e.g. rents) and amenities, thereby forming a job ladder in overall desirability (a "Mortensen motive" for amenity variation). High-SES families may provide advantages in climbing this job ladder, for example through networks, information, or skill-based sorting , giving their children access not only to higher pay but also better amenities and thus reinforcing overall inequality.[33] This joint presence of "Rosen-" and "Mortensen-"type variation complicates efforts to quantify the role of amenities in firm desirability variation, but we provide evidence of a positive SES gradient in amenities *conditional* on firm pay.

## 5.1 Inferring SES gradients in firm values using job-to-job flows

As a starting point, we quantify the extent to which firms "poach" workers from other firms rather than hiring from non-employment – a simple, widely used proxy for overall firm desirability (Bagger and Lentz, 2019; Crane et al., 2023). Figure 9a shows that high-paying firms (with high $\hat{\psi}_j$) predominantly hire workers from other firms, while low-paying firms more often hire from non-employment, where workers typically have weaker outside options. This pattern aligns with the job ladder implied by standard search models (Burdett and Mortensen,

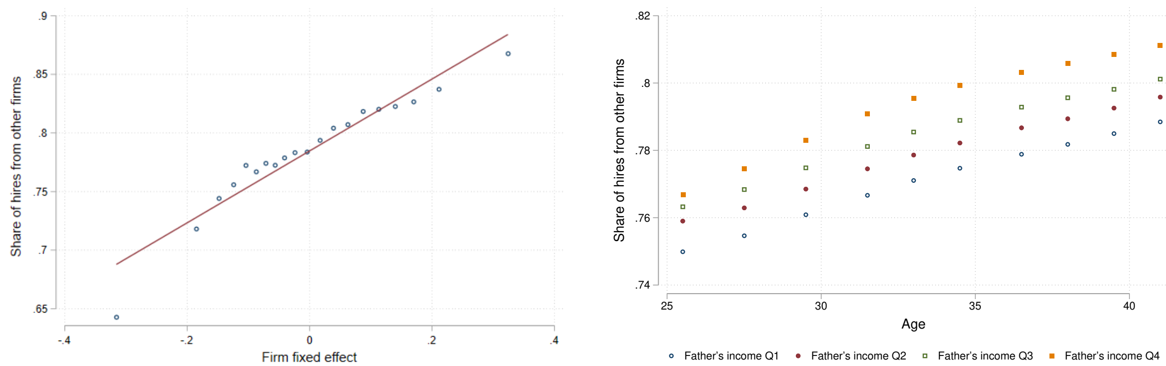
[32]For example, variation in amenities conditional on pay would arise if firms are equally profitable but face varying costs of providing better amenities.

[33]This "Mortensen motive" of amenity variation is not as straightforward to identify in the data (Sorkin, 2018).

Figure 9: Poaching rates from other firms

(a) Poaching rate vs firm pay premium

(b) Poaching rates by father's income




Notes: Subfigure (a) shows a binned scatterplot of the share of hires from employment ("poaching rate") on the estimated AKM firm fixed effects based on equation (1). Subfigure (b) shows the mean poaching rate by age and quartile of father's income.

1998) and suggests that high-paying firms are indeed more attractive to workers.

Figure 9b shows that the poaching rate of the current employer increases with parental income and increases more steeply over age for children from the top quartile of parental income, mirroring the pattern of firm pay premia. However, while average firm pay peaks in the early 30s and then declines slightly (cf. Figure 2), poaching rates continue to rise throughout workers' 30s. Importantly, these SES gaps in poaching rates disappear when conditioning on firm pay (Appendix Figure A19), providing a first indication that pay gaps are reasonable proxies for SES gradients in broader firm value.[34]

Of course, poaching is only a coarse indicator for overall firm desirability. To obtain a more comprehensive measure of firm value, or $V_j$, we implement the revealed-preference approach of Sorkin (2018), which uses all voluntary job-to-job transitions to infer workers' ordinal preferences over firms. We define voluntary employer-to-employer transitions as moves that do not involve unemployment benefit receipt or zero-earnings spells between consecutive firm spells.

Estimated firm values $\hat{V}_j$ are strongly correlated with estimated firm pay premia $\hat{\psi}_j$. In our data, the correlation is 0.52 (Appendix Table A13), very similar to the US estimate in Sorkin (2018). This strong correlation suggests either that pay is a key driver of job-to-job flows or that pay is positively correlated with non-pay amenities. In either case, firm pay premia appear to be informative about overall firm desirability.

To draw more direct inference about differences in welfare, we examine SES gradients in firm value. Table 6 reports how both firm pay and firm value vary with parental income. Panel A presents baseline estimates, while Panels B and C explore the roles of skill sorting and sampling error. Because the firm values lack an absolute scale, we standardize $\hat{V}_j$ to

[34] As shown in Appendix Figure A19, quit rates also increase in parental income, perhaps reflecting that high-SES children switch firms more often (Figure 3a) and work in faster-growing firms (Figure 5b).

Table 6: SES gaps in firm values

| | Firm pay $\hat{\psi}_j$ | | Firm value $\hat{V}_j$ | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Panel A: Baseline | | | |
| $y_{f(i)}$ | 0.061*** | 0.037*** | 0.047*** | 0.016*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\hat{V}_j$ | | 0.516*** | | |
| | | (0.001) | | |
| $\hat{\psi}_j$ | | | | 0.508*** |
| | | | | (0.001) |
| R-squared | 0.031 | 0.285 | 0.019 | 0.276 |
| Observations | 752,061 | 752,061 | 752,061 | 752,061 |
| | Panel B: Controlling for skill sorting | | | |
| $y_{f(i)}$ | 0.044*** | 0.029*** | 0.032*** | 0.010*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\hat{V}_j$ | | 0.481*** | | |
| | | (0.001) | | |
| $\hat{\psi}_j$ | | | | 0.481*** |
| | | | | (0.001) |
| $\hat{\alpha}_i$ | 0.152*** | 0.088*** | 0.133*** | 0.060*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| R-squared | 0.096 | 0.305 | 0.070 | 0.285 |
| Observations | 752,061 | 752,061 | 752,061 | 752,061 |
| | Panel C: Split-sample IV | | | |
| $y_{f(i)}$ | 0.058*** | 0.033*** | 0.045*** | 0.014*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\hat{V}_j$ | | 0.564*** | | |
| | | (0.001) | | |
| $\hat{\psi}_j$ | | | | 0.541*** |
| | | | | (0.001) |
| R-squared | 0.030 | 0.278 | 0.019 | 0.277 |
| Observations | 689,202 | 689,202 | 689,202 | 689,202 |

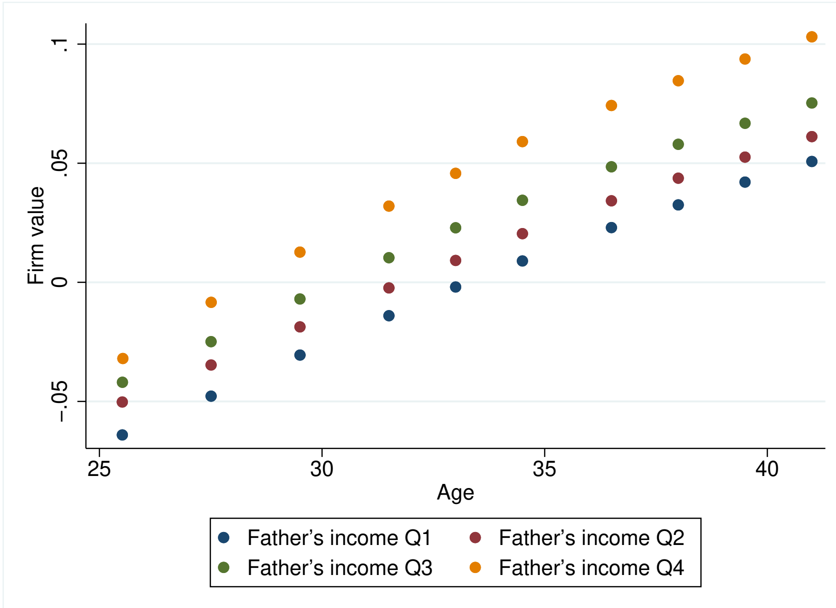
Notes: Column (1) reports the slope coefficient from regressing the estimated firm pay $\hat{\psi}_j$ from eq. (1) on father's log income for the subsample of firms that are included in the model to estimate the firm values. Column (3) reports the slope coefficient from regressing the estimated firm value $\hat{V}_j$ based on Sorkin (2018) on parental income. Columns (2) and (4) repeat these regressions but condition on the estimated firm value or firm pay, respectively. In Panel B, we add controls for the estimated worker fixed effect from eq. (1). In Panel C, we use split-sample techniques to instrument the estimated firm value or pay; see Section 5.2 for details. The Estimated firm values $\hat{V}_j$ are standardized to mean zero and standard deviation one. Robust standard errors in parentheses.

mean zero and to have the same standard deviation as $\hat{\psi}_j$.

Column 1 shows that the SES gradient in firm pay premia (0.061) for the subsample with estimated firm values resembles that of the baseline sample (0.065).[35] Column 3 demonstrates

[35]To estimate the firm value $V_j$, a firm needs to be part of a more restrictive *strongly* connected set in terms of voluntary firm-to-firm transitions (see Sorkin, 2018). The sample size is thus slightly reduced.

Figure 10: Firm values over the lifecycle




Note: The figure plot the firm value over the life cycle, by quartile of father's income. The firm value is estimated following following Sorkin (2018).

that this SES gradient extends to overall firm value, and thus the overall desirability of the firms. Thus, sorting into higher-value employers makes children from high-SES families better off. Although the absolute size of the SES gradients in firm pay and value are not comparable, we find that parental income explains a smaller share of the variation in firm value than in firm pay ($R^2$ of 0.019 vs. 0.031).

Figure 10 shows how this gradient in firm value evolves over the lifecycle. As with firm pay, gaps in firm value emerge at the start of the career and widen with age. However, while firm pay gaps stabilize by the early 30s, SES gaps in firm value continue to grow, as does the average firm value. These patterns suggest that, as careers progress, high-SES workers increasingly sort into firms offering better amenities or other non-pay benefits.

As shown in Panel B of Table 6, the (mid-career) SES gradients in firm pay and firm value drop by about one-third when controlling for individual fixed effects. This indicates that skill differences is a partial explanation to why high-SES workers sort into more desirable firms. Notably, the magnitude of skill sorting is very similar for firm values and firm pay premia (cf. Panels A and B, columns 1 and 3). Yet, even conditional on skills, high-SES children remain considerably more likely to work in higher-value firms.

The firm value may reflect both rents and non-priced amenities, and without further assumptions we cannot distinguish the relative roles of these components. Nor do these results indicate whether high-SES children receive better or worse amenities in absolute terms — only that the *value* they derive from firm pay is larger than any compensating disamenities. However, the positive SES gradient in firm values, combined with the strong correlation between firm pay premia and firm values, suggests that variation in pay premia reflects genuine variation in overall firm desirability (in the "Mortensen" sense) rather than solely compensating differentials.

Next, we examine SES gradients in firm pay conditional on firm value, and vice versa. In column (2), we regress firm pay on both parental income and estimated firm values. The coefficient on parental income is nearly halved compared to column (1), but remains positive. Conditional on working at similarly valued firms, children from high-SES backgrounds thus appear to sort into firms that offer higher pay but lower non-wage amenities (compensating differentials). This interpretation, however, hinges on taking the estimated firm values at face value; as we show below, measurement error would affect these patterns.

Column 4 shows the reverse regression, the SES gradient in firm value conditional on firm pay. Conditioning on firm pay reduces the SES gradient in firm value by about two-thirds (cf. Panels A and B, columns 3 and 4), yet the gradient remains positive. Thus, among firms offering similar pay premia, children from high-SES families sort into more desirable firms than their low-SES peers.[36]

At first glance, the estimates in columns (2) and (4) may seem contradictory: high-SES children are in firms that are more desirable than their pay would imply ("reverse regression", column 4), but also in firms that pay more than their firm value implies ("forward regression", column 2). This pattern, however, is consistent with differences in both compensating differentials and rents. For instance, it would arise if high-SES children work in firms offering both high pay *and* high value, while low-SES children sort into firms offering either high pay *or* high value, but rarely both.[37]

Table 6 thus provides several insights. First, column (2) suggests that, among workers in equally valued firms, high-SES children are in firms offering relatively more pay and worse non-wage amenities, consistent with a stronger taste for pay. Second, column (4) shows that, among workers with equal firm pay, high-SES children work in firms offering higher overall utility. Since firm pay consists of both rents and (negative) amenity compensation, and only the former contributes to the firm value, the results suggest that high pay tends to reflect compensating differentials for low-SES children, but rents for high-SES children.[38]

While these results are suggestive of high-SES children placing a strong value on pay, we are unable to conclude whether the overall SES gradient in amenities is positive or negative, given the possibility of Mortensen-type variation in amenities (unconditional on firm values).

[36] When we additionally control for individual fixed effects (Panel B), the SES gradient declines by about one-third, indicating that skill sorting partially explains why high-SES children are more likely to enter high-value firms, conditional on firm pay.

[37] See also Goldberger (1984), who explains why forward and reverse regressions can yield seemingly conflicting results in studies of labor market discrimination. While forward and reverse regressions give the same qualitative result in deterministic settings, they can diverge when the underlying relationships are stochastic.

[38] This finding is in line with evidence by Haeck and Laliberté (2025), who show that low-SES children are overrepresented in physically demanding industries such as mining, oil and gas, or construction.

## 5.2 Firm pay, firm values, and measurement error

An alternative interpretation of the patterns in Table 6 is that our estimates of firm pay and firm value are affected by measurement error. Estimating these objects from job-to-job flows for each firm is demanding and subject to sampling noise, particularly for smaller firms with few observed movers. While this issue is mitigated by the use of population-wide data and by excluding firms with very few movers, some sampling variability inevitably remains. Such noise inflates the variances of both $\hat{\psi}$ and $\hat{V}$. This is less concerning when firm pay and value are used as dependent variables, as in columns (1) and (3) of Table 6, but would introduce an attenuation bias when using them as controls on the right-hand side of a regression – which would then spill over to the coefficient on parental income.

To illustrate why measurement error would induce a *positive* bias in the coefficient on parental income, assume that both the estimated firm value $\hat{V}$ and estimated firm pay $\hat{\psi}$ are noisy proxies for their true counterparts,

$$\hat{V}_j = V_j + u_j^V \tag{20}$$

$$\hat{\psi}_j = \psi_j + u_j^\psi \tag{21}$$

where $u_j^V$ and $u_j^\psi$ are classical measurement error with variances $\sigma^2_{u^V}$ and $\sigma^2_{u^\psi}$, respectively. Moreover, assume that firm value $V_j$ is the sum of firm pay $\psi_j$ and non-pay amenities $a_j$,

$$V_j = \psi_j + a_j, \tag{22}$$

and that firm pay increases in parental income according to

$$\psi_j = \gamma y_f + \varepsilon_j \tag{23}$$

with $\gamma > 0$ and $\varepsilon_j$ capturing heterogeneity in firm pay orthogonal to $y_f$, with variance $\sigma^2_\varepsilon$.

In this model, the coefficient $\beta_y$ on parental income $y_f$ in a regression of estimated firm value $\hat{V}_j$ on estimated pay premia $\hat{\psi}_j$ and parental income $y_f$ – corresponding to column (4) of Table 6 – is positive even when amenities $a_j$ are uncorrelated with parental income or firm pay. Specifically, by the Frisch-Waugh-Lovell theorem (see Appendix A6.1), this coefficient converges in probability to

$$\beta_y = \gamma \frac{\sigma^2_{u^\psi}}{\sigma^2_\varepsilon + \sigma^2_{u^\psi}}, \tag{24}$$

which is positive if $\sigma^2_{u^\psi} > 0$, that is, whenever estimated firm pay premia $\hat{\psi}_j$ are an imperfect proxy for true firm pay $\psi_j$. Thus, Measurement error in firm pay pushes the conditional SES gradient in firm value upwards. An analogous argument applies to the reverse regression of estimated firm pay $\hat{\psi}_j$ on $\hat{V}_j$ and $y_f$.

To assess the role of sampling error empirically, we implement a split-sample instrumental variables (IV) approach, as for example in Sorkin (2018) and Drenik et al. (2023). We randomly split individuals into two equally sized groups, estimate both firm pay premia and firm values in each subsample, and use the estimate from one subsample as an instrument for the corresponding estimate from the other. Because each subsample must contain enough movers per firm, this procedure reduces the number of firms and individual observations relative to the baseline.

Panel C of Table 6 presents the results. When firm pay or firm value is the dependent variable (Columns 1 and 3), the estimates remain similar to the baseline. However, when using the split-sample IV strategy to instrument the estimated firm value $\hat{V}_j$ (Column 3) and firm pay $\hat{\psi}_j$ (Column 4), we observe an increase in the respective coefficients and a corresponding decrease in the coefficient on parental income. This pattern is consistent with classical measurement error in the firm-level regressors, which the IV procedure helps correct. However, the magnitude of this adjustment is relatively small – around 10 percent – and the SES gradients remain positive and substantial.

We therefore conclude that sampling noise is not a major driver of the conditional SES gradients in Table 6. This finding aligns with Sorkin (2018), who also reports limited influence of sampling variability in a similar setting. Of course, other forms of measurement error – perhaps stemming from conceptual limitations or model misspecification – may still be present. Nevertheless, the evidence suggests that firm pay is a useful proxy for overall firm desirability, and that the observed SES gradient in firm pay translates into a corresponding gradient in broader firm value, only part of which is attributable to 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 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.

Altonji, J. and C. Pierret (2001): "Employer Learning and Statistical Discrimination*," *Quarterly Journal of Economics*, 116, 313–350.

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," .

Bagger, J. and R. Lentz (2019): "An Empirical Model of Wage Dispersion With Sorting," *The Review of Economic Studies*, 86, 153–190.

Barrios-Fernández, A., C. Neilson, and S. Zimmerman (2024): "Elite Universities and the Intergenerational Transmission of Human and Social Capital," Tech. rep., Working Paper, Universidad de Los Andes.

Battiston, D., M. Friedrich, and I. Helm (2024): "Workplace Heterogeneity in Wage Growth," Tech. rep., Working Paper, University Linz.

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.

BONHOMME, S., T. LAMADON, AND E. MANRESA (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–739.

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.

CAMPA, P. (2025): "No Kid Is an Island: Intergenerational Mobility and Peer Effects," Tech. rep., Working Paper, University of Geneva.

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.

CARD, D., J. ROTHSTEIN, AND M. YI (2024): "Industry wage differentials: A firm-based approach," *Journal of Labor Economics*, 42, S11–S59.

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

CRANE, L. D., H. R. HYATT, AND S. M. MURRAY (2023): "Cyclical Labor Market Sorting," *Journal of Econometrics*, 233, 524–543.

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," .

DRENIK, A., S. JÄGER, P. PLOTKIN, AND B. SCHOEFER (2023): "Paying Outsourced Labor: Direct Evidence from Linked Temp Agency-Worker-Client Data," *The Review of Economics and Statistics*, 105, 206–216.

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

ELIASON, M., L. HENSVIK, F. KRAMARZ, AND O. N. SKANS (2023): "Social connections and the sorting of workers to firms," *Journal of Econometrics*, 233, 468–506.

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 (2025): "Firms and the Intergenerational Transmission of Labor Market Advantage," *American Journal of Sociology*, 131, 322–370.

FARBER, H. S. AND R. GIBBONS (1996): "Learning and Wage Dynamics," *The Quarterly Journal of Economics*, 111, 1007–1047.

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.

GIBBONS, R., L. F. KATZ, T. LEMIEUX, AND D. PARENT (2005): "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 23, 681–724.

GOLDBERGER, A. S. (1984): "Reverse Regression and Salary Discrimination," *Journal of Human Resources*, 293–318.

GOLDTHORPE, J. H. (2014): "The Role of Education in Intergenerational Social Mobility: Problems from Empirical Research in Sociology and Some Theoretical Pointers from Economics," *Rationality and Society*, 26, 265–289.

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

HAECK, C. AND J.-W. LALIBERTÉ (2025): "Careers and Intergenerational Income Mobility," *American Economic Journal: Applied Economics*, 17, 431–58.

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. (2024): "Firm Wage Effects," *Handbook of Labor Economics*, 5, 115–181.

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, R. SAGGIO, AND S. A. WOODBURY (2023): "Do Firm Effects Drift? Evidence from Washington Administrative Data," *Journal of Econometrics*, 233, 375–395.

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

LALIBERTÉ, J.-W. AND A. WHALLEY (2025): "Social Connections and the Persistence of Income Across Generations," Tech. rep., Working Paper, University of Calgary.

LAMADON, T., M. MOGSTAD, AND B. SETZLER (2022): "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, 112, 169–212.

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.

MICHELMAN, V., J. PRICE, AND S. D. ZIMMERMAN (2022): "Old Boys' Clubs and Upward Mobility Among the Educational Elite," *Quarterly Journal of Economics*, 137, 845–909.

MORCHIO, I. AND C. MOSER (2024): "The Gender Pay Gap: Micro Sources and Macro Consequences," Tech. rep., National Bureau of Economic Research.

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

——— (2024): "Intergenerational income mobility," in *Research Handbook on Intergenerational Inequality*, ed. by J. Blanden, J. Erola, E. Kilpi-Jakonen, and L. Macmillan, Edward Elgar Publishing, 56–72.

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

PIOPIUNIK, M., G. SCHWERDT, L. SIMON, AND L. WOESSMANN (2020): "Skills, signals, and employability: An experimental investigation," *European Economic Review*, 123, 103374.

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.

SCHÜLE, P. (2026): "Career Preferences and Socio-economic Background," *Labour Economics*, 98, 102814.

SOCKIN, J. (2022): "Show Me the Amenity: Are Higher-paying Firms Better All Around?" Tech. rep., CESifo Working Paper.

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 | 0.268 | 0.242 | 0.205 |
| Components: | | | |
| Individual FEs | 0.101 (37.7%) | 0.072 (29.8%) | 0.073 (35.6%) |
| Firm FEs | 0.022 (8.2%) | 0.026 (10.7%) | 0.026 (12.7%) |
| Covariance (sorting) | 0.019 (7.1%) | 0.022 (9.1%) | 0.024 (11.7%) |
| Covariates and residual | 0.125 (46.6%) | 0.121 (50.0%) | 0.082 (40.0%) |
| Worker obs. | 7,482,143 | 960,925 | 836,743 |
| Number of firms | 460,875 | 314,598 | 172,225 |
| Worker-year obs. | 118,832,220 | 12,977,062 | 2,352,856 |

Notes: Variance decomposition of log earnings into the components of equation (1). Column (1) shows the variance decomposition for the AKM sample, column (2) for the main lifecycle sample, and column (3) for the main sample with mean earnings estimated for 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. Firm fixed effects explain between 7-11% of the variance in log earnings, while the covariance between firm and worker fixed effects explain another 7-11%. We therefore find strong sorting between high-wage workers and firms: in the AKM sample, the implied correlation coefficient between the estimated worker and firm fixed effects is 0.42 $(0.019/(\sqrt{0.103}\sqrt{0.020}))$.

Differences in firm pay and worker-firm sorting together explain 14-22% of the total variance, depending on sample. 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 Alternative measures of parental income

Table A2: Decomposition of the IGE using average parental income

| | 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) |
| $y_{f(i)}$ | 0.323*** | 0.179*** | 0.103*** | 0.040*** | 0.076*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| $\hat{\alpha}_i$ | | | | | 0.152*** |
| | | | | | (0.001) |
| Share of IGE | 100% | 55% | 32% | 0.12% | 24% |
| Worker obs. | 777,364 | 777,364 | 777,364 | 777,364 | 777,364 |

Notes: Column (1) reports the estimated slope coefficient from regression (2) of mean of log child earnings at ages 39-41 on the average of father's and mother'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 $\alpha_i$, mean firm fixed effects $\psi_j$ for age 39-41, and time-varying controls. Robust standard errors in parentheses.

### A1.2.2 Measurement error

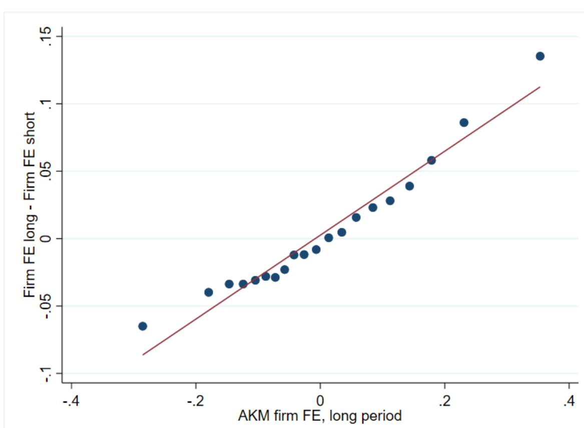
The table shows the IGE decomposition when the AKM model have been estimated for shorter time period, for the years 2010-2015. Mean earnings and mean firm fixed effect are then calculated for all these years.

Table A3: Decomposition of the IGE using a shorter AKM period

| | 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 | | | | |
| $y_{f(i)}$ | 0.193*** | 0.157*** | 0.038*** | 0.030*** |
| | (0.001) | (0.001) | (0.000) | (0.000) |
| $\hat{\alpha}_i$ | | | | 0.056*** |
| | | | | (0.000) |
| Share of IGE | 1.0 | 0.81 | 0.20 | 0.16 |
| Worker obs. | 749,055 | 749,055 | 749,055 | 749,055 |
| B: AKM estimated for the years 1985-2018 (intergenerational sample same as in panel A) | | | | |
| $y_{f(i)}$ | 0.193*** | 0.115*** | 0.061*** | 0.043*** |
| | (0.001) | (0.001) | (0.000) | (0.000) |
| $\hat{\alpha}_i$ | | | | 0.160*** |
| | | | | (0.001) |
| Share of IGE | 1.0 | 0.60 | 0.31 | 0.22 |
| Worker obs. | 749,055 | 749,055 | 749,055 | 749,055 |

Notes: Estimated slope coefficients from regression (2) of mean of log child earnings over the ages 39-41 on log father's earnings and corresponding regressions when decomposing child log earnings $y_{ijt}$ according to equation (1) into individual fixed effects $\alpha_i$ and mean firm fixed effects $\psi_j$ over the ages 39-41. The columns for time-varying control is not included since we have imputed firm values for the ages 39-41, 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 between AKM firm premia for short and long period




Notes: Binned scatter plot of the difference between the estimated AKM firm premia for the long period (1985-2018) and the short period (2010-2015) against the firm premia for the long period.

Table A4: Decomposition of IGE using a random AKM subsample

| | 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 | | | | |
| $y_{f(i)}$ | 0.197*** | 0.118*** | 0.062*** | 0.044*** |
| | (0.002) | (0.001) | (0.001) | (0.001) |
| $\hat{\alpha}_i$ | | | | 0.150*** |
| | | | | (0.001) |
| Share of IGE | 1.0 | 0.60 | 0.31 | 0.22 |
| Worker obs. | 407,750 | 407,750 | 407,750 | 407,750 |
| B: AKM estimated for full sample (intergenerational sample same as in panel A) | | | | |
| $y_{f(i)}$ | 0.197*** | 0.117*** | 0.062*** | 0.044*** |
| | (0.002) | (0.001) | (0.001) | (0.001) |
| $\hat{\alpha}_i$ | | | | 0.159*** |
| | | | | (0.001) |
| Share of IGE | 1.0 | 0.59 | 0.31 | 0.22 |
| Worker obs | 407,750 | 407,750 | 407,750 | 407,750 |

Notes: Estimated slope coefficients from regression (2) of mean of log child earnings over the ages 39-41 on log father's earnings and corresponding regressions when decomposing child log earnings $y_{ijt}$ according to equation (1) into individual fixed effects $\alpha_i$ and mean firm fixed effects $\psi_j$ over the ages 39-41. 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.

#### A1.2.3 Alternative specifications for the AKM regression

Table A5: Decomposition of the IGE using alternative AKM specifications

| | 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. AKM estimated with time-varying firm FE | | | | |
| $y_{f(i)}$ | 0.198*** | 0.112*** | 0.071*** | 0.015*** | 0.053*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.57 | 0.36 | 0.076 | 0.27 |
| Worker obs. | 826,314 | 826,314 | 826,314 | 826,314 | 826,314 |
| | B. AKM estimated using establishment codes | | | | |
| $y_{f(i)}$ | 0.198*** | 0.113*** | 0.068*** | 0.018*** | 0.050*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.57 | 0.34 | 0.09 | 0.25 |
| Worker obs. | 831,883 | 831,883 | 831,883 | 831,883 | 831,883 |
| | C. AKM estimated using firm codes for the largest firms | | | | |
| $y_{f(i)}$ | 0.198*** | 0.128*** | 0.052*** | 0.019*** | 0.033*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.65 | 0.26 | 0.10 | 0.16 |
| Worker obs. | 836,553 | 836,553 | 836,553 | 836,553 | 836,553 |
| | D. AKM estimated without excluding firms with few movers | | | | |
| $y_{f(i)}$ | 0.197*** | 0.112*** | 0.069*** | 0.017*** | 0.054*** |
| | (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. | 885,220 | 885,220 | 885,220 | 885,220 | 885,220 |
| | E. Excluding firms with less than 10 movers | | | | |
| $y_{f(i)}$ | 0.197*** | 0.117*** | 0.062*** | 0.018*** | 0.044*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.59 | 0.31 | 0.09 | 0.22 |
| Worker obs. | 814,057 | 814,057 | 814,057 | 814,057 | 814,057 |
| | F. Excluding firms with less than 50 movers | | | | |
| $y_{f(i)}$ | 0.198*** | 0.123*** | 0.056*** | 0.019*** | 0.038*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.62 | 0.28 | 0.1 | 0.19 |
| Worker obs. | 696,157 | 696,157 | 696,157 | 696,157 | 696,157 |

Notes: Estimated slope coefficients from regression (2) of mean of log child earnings at ages 39-41 on log father's earnings and corresponding regressions when decomposing child earnings $y_{ijt}$ according to equation (1). The different panels show different variants of estimating equation (1). In panel A, we estimate time-varying firm fixed effects by dividing the period 1985-2018 into four 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, panel E excludes firms with less than 10 movers, and panel F excludes firms with less than 50 movers (baseline: at least 5 movers).

Table A6: Decomposition of the IGE using AKM estimated with wages

| | Dependent variable | | | | |
|---|---|---|---|---|---|
| | $y_{ijt}$ | $\hat{\alpha}_i$ | $\hat{\psi}_{j=J(i,t)}$ | $\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$ | $\hat{\psi}_{j=J(i,t)}$ |
| | (1) | (2) | (3) | (4) | (5) |
| | A: AKM estimated with wages | | | | |
| $y_{f(i)}$ | 0.174*** | 0.149*** | 0.029*** | -0.004*** | 0.017*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| $\hat{\alpha}_i$ | | | | | 0.086*** |
| | | | | | (0.000) |
| Share of IGE | 1.0 | 0.86 | 0.16 | 0.02 | 0.10 |
| Worker obs. | 541,778 | 541,778 | 541,778 | 541,778 | 541,778 |
| | B: AKM estimated with earnings (excluding individuals not in wage sample) | | | | |
| $y_{f(i)}$ | 0.191*** | 0.116*** | 0.055*** | 0.002*** | 0.036*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| $\hat{\alpha}_i$ | | | | | 0.163*** |
| | | | | | (0.001) |
| Share of IGE | 1.0 | 0.61 | 0.28 | 0.01 | 0.19 |
| Worker obs. | 541,778 | 541,778 | 541,778 | 541,778 | 541,778 |

Notes: Estimated slope coefficients from regression (2) of mean of log child wages or earnings at ages 39-41 on log father's earnings and corresponding regressions when decomposing child wage $y_{ijt}$ according to equation (1). 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.

### A1.2.4 Decomposition of the IGE: Heterogeneity

Table A7 shows the decomposition of the intergenerational earnings elasticity for different subsamples.

Table A7: Heterogeneity in the IGE decomposition

| | Dependent variable | | | | |
|---|---|---|---|---|---|
| | $y_{ijt}$ | $\hat{\alpha}_i$ | $\hat{\psi}_{j=J(i,t)}$ | $\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}}+\hat{\varepsilon}_{ijt}$ | $\hat{\psi}_{j=J(i,t)}\lvert\hat{\alpha}_i$ |
| | (1) | (2) | (3) | (4) | (5) |
| | A. Sample: Men | | | | |
| $y_{f(i)}$ | 0.229*** | 0.138*** | 0.070*** | 0.021*** | 0.050*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Share of IGE | 1.0 | 0.60 | 0.31 | 0.09 | 0.22 |
| Worker obs. | 428,245 | 428,245 | 428,245 | 428,245 | 428,245 |
| | B. Sample: Women | | | | |
| $y_{f(i)}$ | 0.165*** | 0.092*** | 0.059*** | 0.014*** | 0.049*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Share of IGE | 1.0 | 0.57 | 0.30 | 0.13 | 0.24 |
| Worker obs. | 408,498 | 408,498 | 408,498 | 408,498 | 408,498 |
| | C. Excluding workers who work in same firm as father | | | | |
| $y_{f(i)}$ | 0.196*** | 0.115*** | 0.063*** | 0.018*** | 0.045*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.000) |
| Share of IGE | 1.0 | 0.59 | 0.32 | 0.09 | 0.23 |
| Worker obs. | 793,295 | 793,295 | 793,295 | 793,295 | 793,295 |
| | D. Excluding public sector firms | | | | |
| $y_{f(i)}$ | 0.227*** | 0.123*** | 0.082*** | 0.022*** | 0.063*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Share of IGE | 1.0 | 0.54 | 0.36 | 0.10 | 0.28 |
| Worker obs. | 536,774 | 536,774 | 536,774 | 536,774 | 536,774 |

Notes: Estimated slope coefficients from regression (2) of mean of log child earnings at ages 39-41 on log father's earnings and corresponding regressions when decomposing child earnings $y_{ijt}$ according to equation (1) into individual fixed effects $\alpha_i$, mean of firm fixed effects $\psi_j$ 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.

#### A1.2.5 Adding controls

Table A8: Decomposing the relationship between firm premia and parental income (occupation sample)

| | Dependent variable: Estimated firm pay premium $\hat{\psi}_{j=J(i,t)}$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| $y_{f(i)}$ | 0.055*** | 0.036*** | 0.035*** | 0.049*** | 0.043*** | 0.032*** | 0.010*** |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) | (0.000) |
| Region FEs | | X | | | | | X |
| Industry FEs | | | X | | | | X |
| Public sector | | | | X | | | X |
| Est. size (log) | | | | | X | | X |
| Occ. FEs | | | | | | X | X |
| Share of $\beta_{firm}$ | 100% | 65% | 64% | 89% | 78% | 58% | 18% |
| Worker obs. | 287,803 | 287,803 | 287,803 | 287,803 | 287,803 | 287,803 | 287,803 |

Notes: Column (1) reports estimates of the slope coefficient from regressing $\hat{\psi}_j$, at age 40 as estimated from equation (1), on log fathers earnings, limiting the sample to individuals with occupation information. Columns (2)-(7) report coefficient estimates from the same regression controlling for region fixed effects (21 counties), industry fixed effects (2-digit level, 59 industries), working in the public sector (public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99), log establishment size, occupation fixed effects (2-digit level), or adding all controls simultaneously. Robust standard errors in parentheses.

## 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 A9 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 $\phi_{akm}$). 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 A10 reports estimates regression (7), showing that the coefficients change only marginally when excluding parental income from the regression.

Table A9: 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 in equations (6a)-(7), regressing each of the skill measures on father's income.

Table A10: Firm pay premia and skills

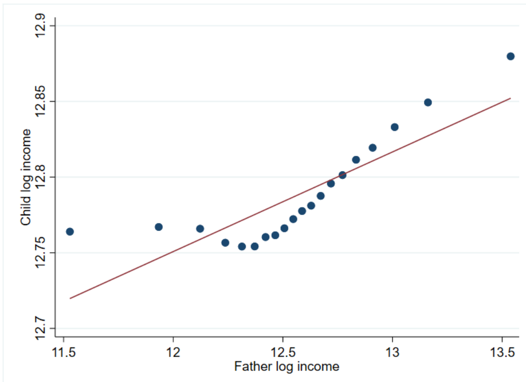
| | Dependent variable | |
|---|---|---|
| | $\hat{\psi}_{j=J(i,t)}$ | $\hat{\psi}_{j=J(i,t)}$ |
| $y_{f(i)}$ | 0.029*** | |
| | (0.001) | |
| Cognitive skills | 0.010*** | 0.011*** |
| | (0.000) | (0.000) |
| Social skills | 0.002*** | 0.002*** |
| | (0.000) | (0.000) |
| Education | 0.003*** | 0.004*** |
| | (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 in equation (5b), regressing the firm premia on father's income and each of the skill measures.

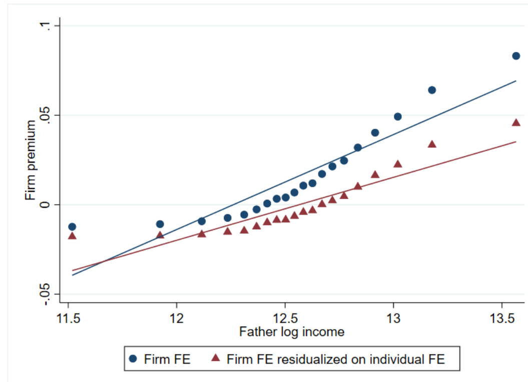
## A2.1 Non-linear firm pay gradients

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

(a) Child vs father's log income



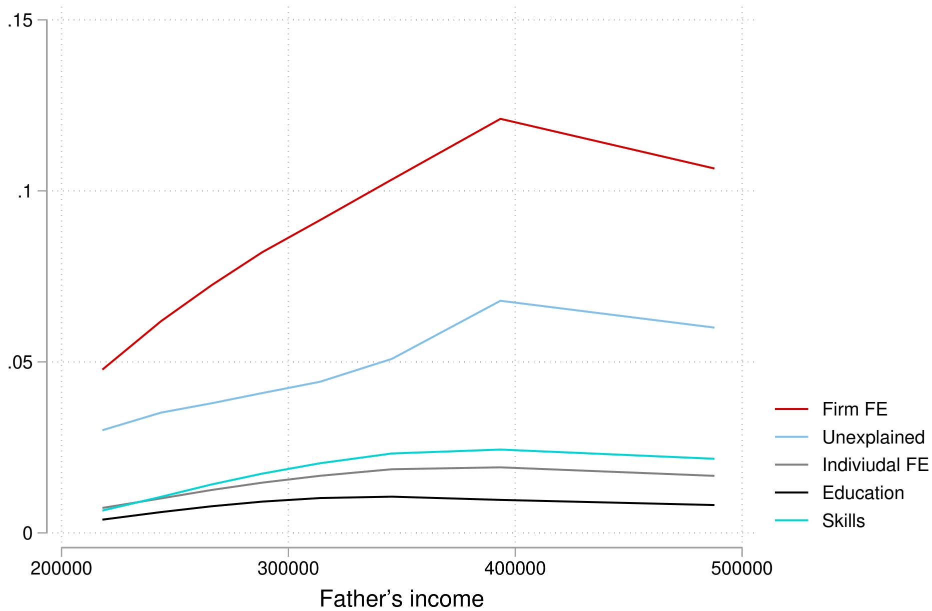

(b) Firm premium at age 40 vs father's log income




Notes: Subfigure (a) shows binned scatter plots of child's log income at age 40 by father's log income. Subfigure (b) shows estimated firm premia $\psi_j$ at age 40 based on equation (1) and firm premia residualized on individual fixed effect by father's log income.

Figures A3 examines how the importance of skill sorting varies across the parental income distribution. Specifically, we follow Hjorth-Trolle and Landersø (2023) and run local-linear regressions of equations (5a)-(7) at each decile of parental income. Figure A3 shows that the firm pay gradient increases markedly along the parental-income distribution, reaching about 0.12 at high levels of parental income. The unexplained part not attributable to skill sorting also grows in absolute magnitude along the distribution, though it declines as a share of $\beta_{firm}$. Both skills (driven by the cognitive measure) and the individual fixed effect grow in importance with parental income, accounting for a substantial share of the raw relationship between firm premia and parental income in the upper-half of the distribution.

Figure A3: Local linear decomposition of skill sorting over the distribution

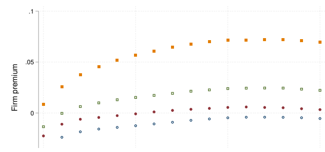



Notes: The figure shows local linear regressions, from the decomposition in equations (5a)-(7) using the main sample, but excluding women and men with missing enlistment scores. Local linear regression around the mean income at each decile of parental income, using a bandwidth of 50,000 Swedish kronor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

# A3 Additional evidence on lifecycle dynamics

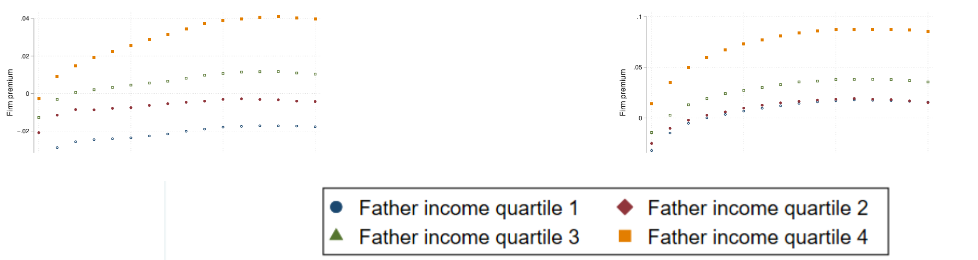
Figure A4: Firm earnings premium over the lifecycle using potential experience

(a) Full sample



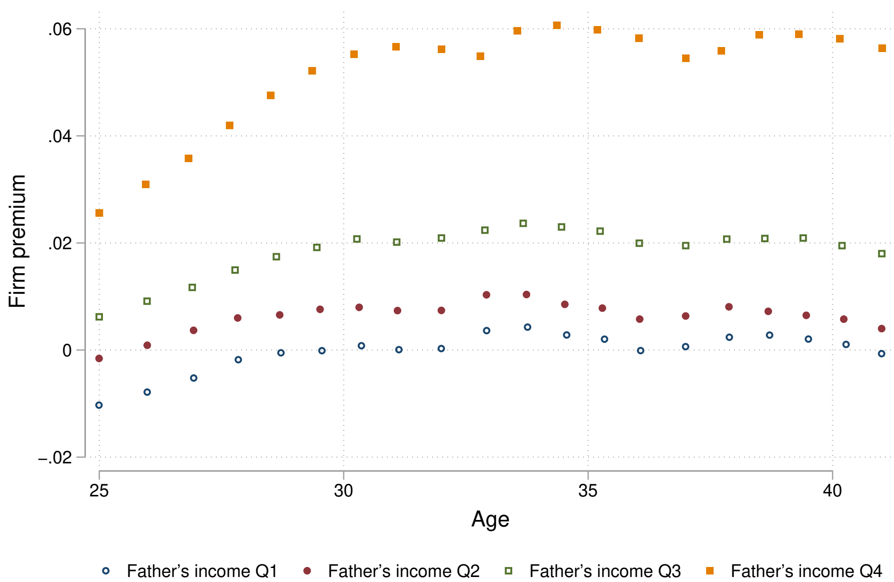

(b) Non-college

(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)). Subfigure (a) shows the result for the full sample, subfigure (b) shows the result for children without collage education and subfigure (c) shows the result for children with college education. Father's income quartiles are defined in the full sample.

Figure A5: Firm pay premium over the lifecycle conditional on individual fixed effects

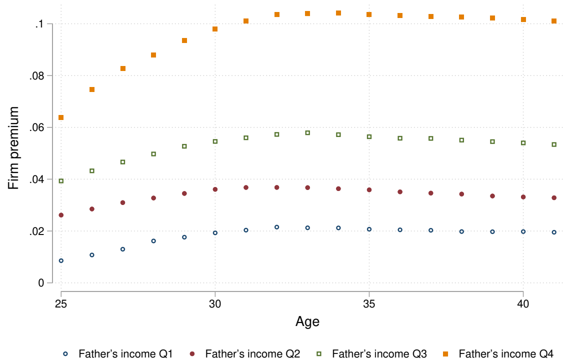




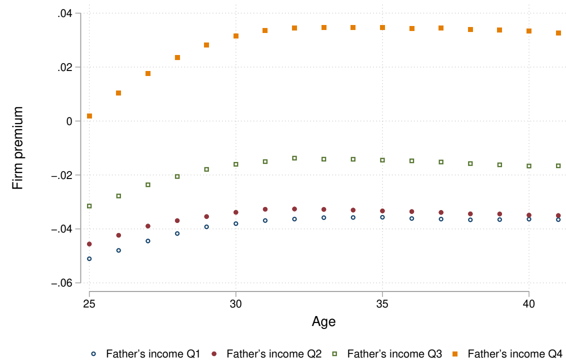

Notes: The figures show the estimated firm premium $\hat{\psi}_j$ from equation (1), residualized on the individual fixed effects, over the life cycle

## Figure A6: Firm pay premium over the lifecycle by gender

(a) Men



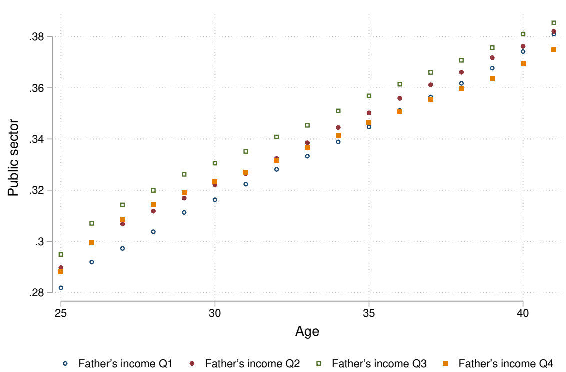

(b) Women



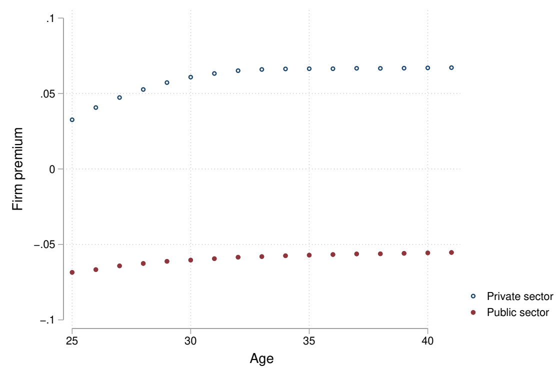

Notes: The figures show the estimated firm premium $\hat{\psi}_j$ from equation (1) over the life cycle, by quartile of father's income. Subfigure (a) shows the results for men and (b) shows the results for women. Fathers' income quartiles are defined in the full sample.

## Figure A7: Public sector

(a) Probability to work in public sector



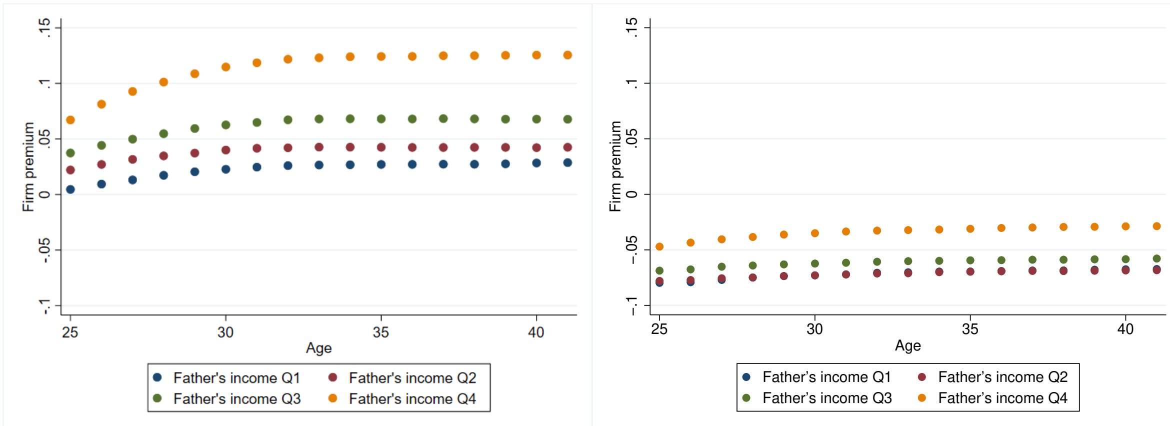

(b) Firm premium




Notes: Figure (a) shows the probability to work in the public sector over the lifecycle by father's income quartile. Figure (b) shows the firm premium separately for private and public sector over the lifecycle. Public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99.

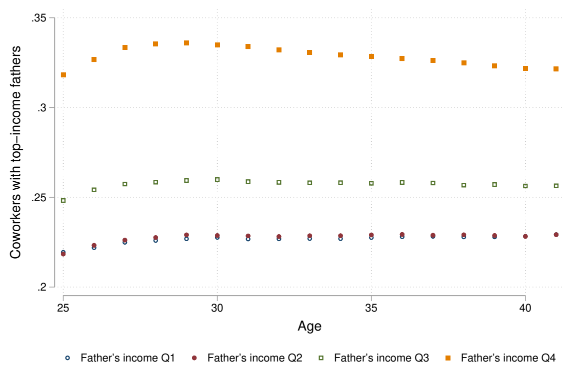
## Figure A8: Firm premia over the lifecycle by public sector

(a) Private sector (b) Public sector




Notes: The figures show the estimated firm premium $\hat{\psi}_j$ from equation (1) over the life cycle, by quartile of father's income. Subfigure (a) shows the results for the private sector and subfigure (b) shows the result for the public sector. Public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99.

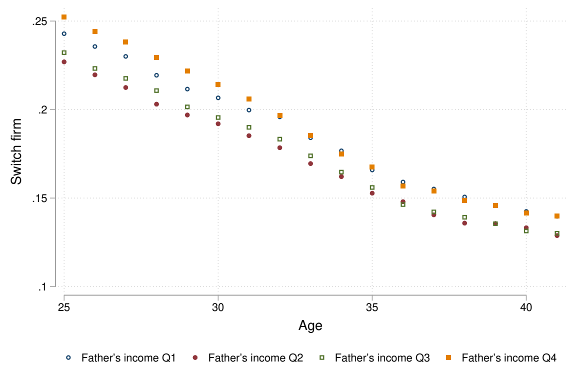
## Figure A9: Share of co-workers with father in top quartile



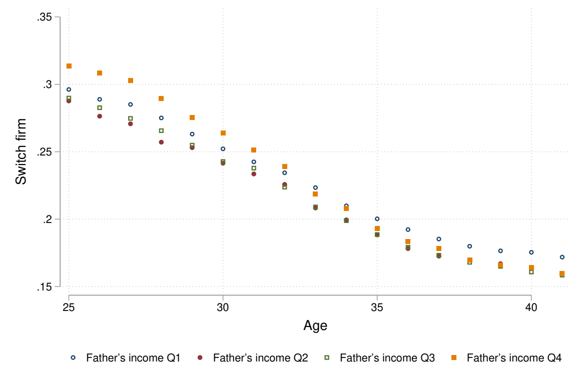
## A3.1 Firm switching

Figure A10: Firm switches by education

(a) Non-college



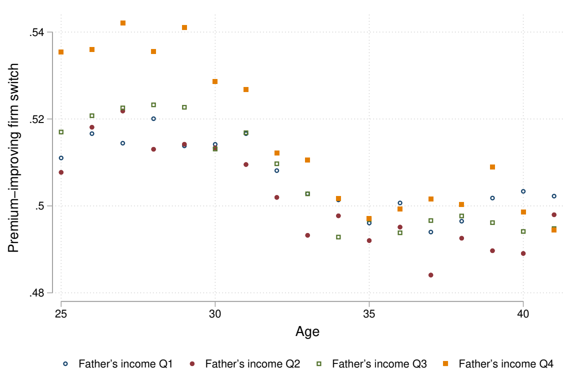

(b) College



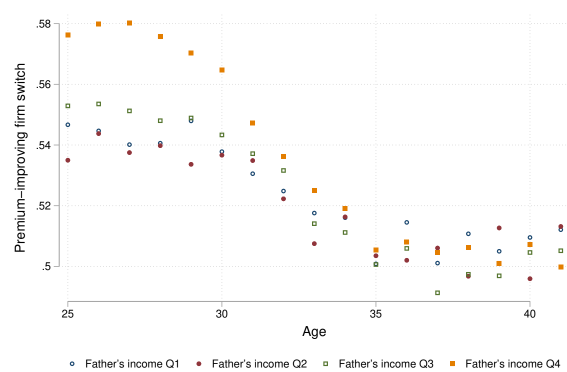

Notes: The figures show pattern for switching firms over the lifecycle by quartile of father's income. Subfigure (a) shows the result for individuals who do not have a college education and subfigure (b) shows the results for individuals who have a college education.

Figure A11: Proportion of premium-improving firm switches

(a) Non-college



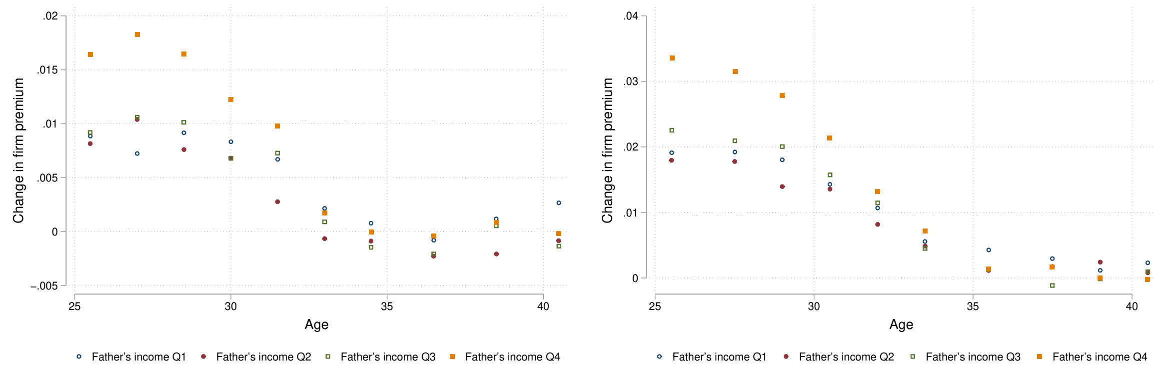

(b) College




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. Subfigure (a) shows the result for individuals who do not have a college education and subfigure (b) shows the results for individuals who have a college education.

Figure A12: Mean change in firm premium among switchers

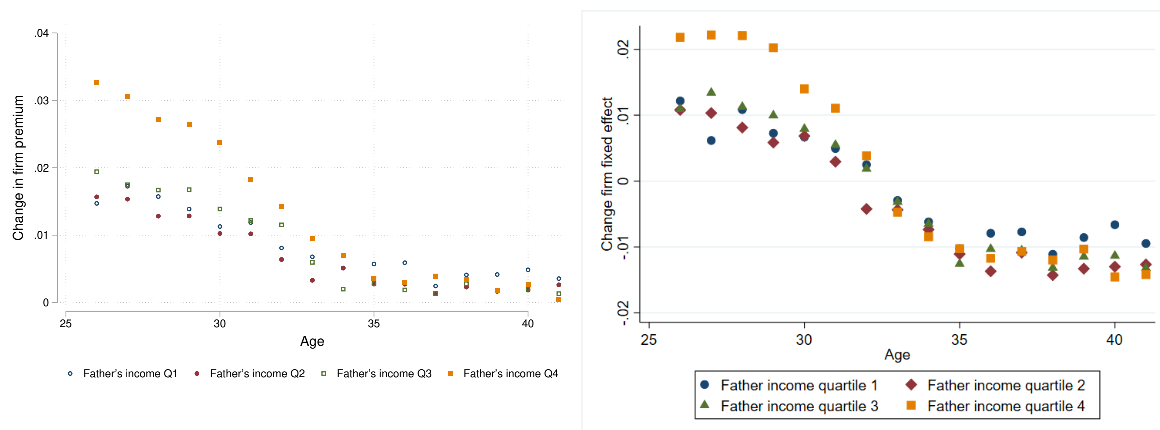
(a) Non-college (b) College




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

Figure A13: 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. Subfigure (a) shows the results for voluntary switches, defined as a switch without any unemployment insurance or without any year with zero income. Subfigure (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 A14 plots the share of individuals who work in the same firm as their father, separately for those with and without a college degree. Children are more likely to work in the same firm as their father at the beginning of their career. Moreover, their probability to work in the same firm is much higher for children from high-earnings fathers than for those from low-income families.

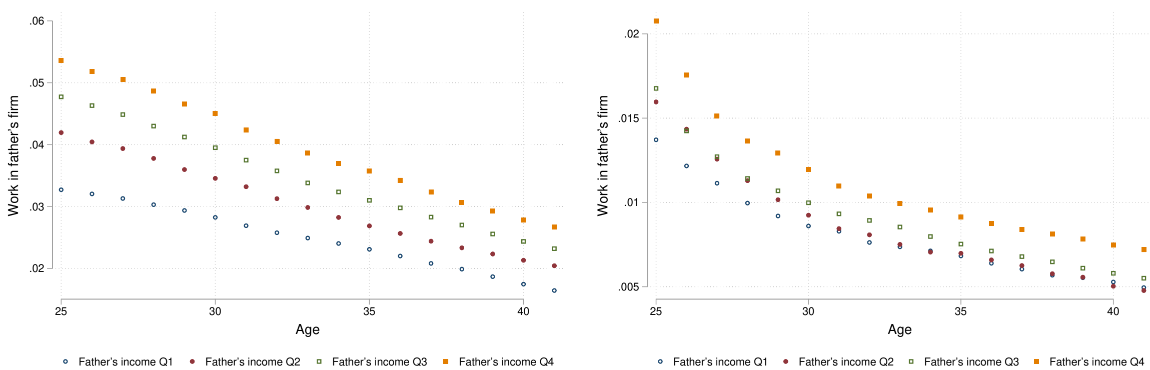
Figure A15 plots the share of individuals who commute (i.e., work and reside in different municipalities), separately for those with and without a college degree. Children from high-income parents are more likely to commute, even within education group, validating the pattern from Figure 4a.

Finally, Table A11 provides event-study type regression results conditional on individual fixed effects, showing how firm pay changes for individuals who begin to commute. While the firm premium increases when individuals start to commute, this increase is much smaller than the raw wage gap between commuters and non-commuters in the cross-section as shown in Figure 4. As shown in column (1), commuting raises firm premia by about 1 percentage points. 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 A14: Working in the same firm as father

(a) Non-college

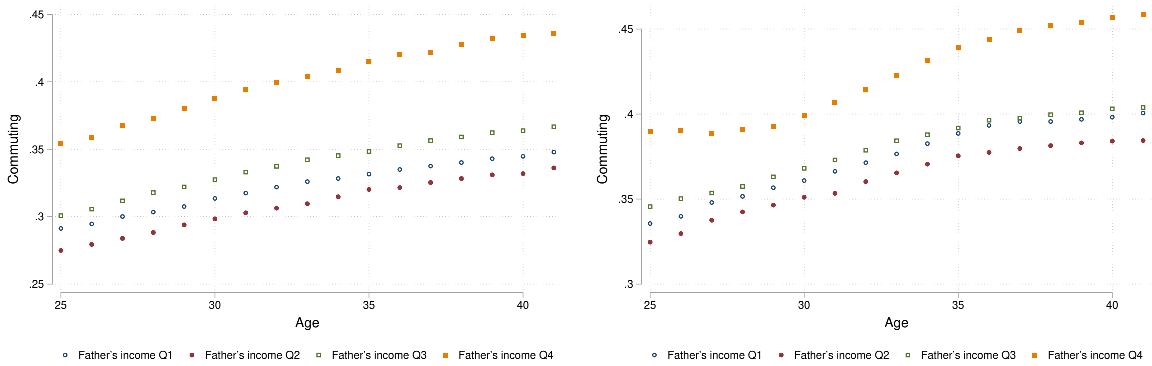
(b) College




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. Subfigure (a) include children without college education and subfigure (b) include children with college education.

## Figure A15: 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 lifecycle, by quartile of father's income. Subfigure (a) includes children without college education and subfigure (b) children with college education.

## Table A11: Commuting

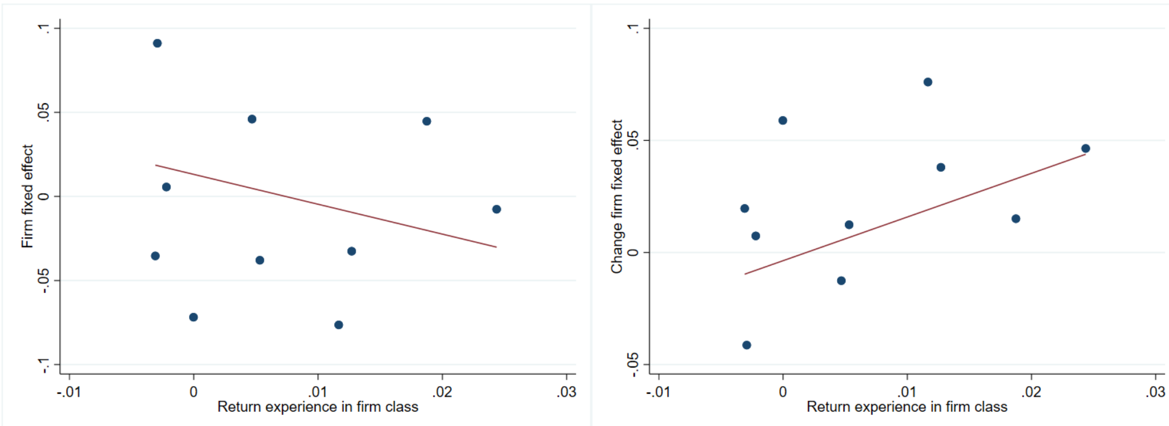
| | Dependent variable: $\hat{\psi}_{j=J(i,t)}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Commuting indicator | 0.010*** | 0.017*** | 0.011*** |
| | (0.000) | (0.000) | (0.000) |
| Commuting # 2nd father's income quartile | | | -0.001** |
| | | | (0.000) |
| Commuting # 3rd father's income quartile | | | 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:: Regression of estimated firm pay premia on an indicator for commuting (working and residing in different municipalities), individual fixed effects, and controls. 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 interacts the commuting indicator with age. Column (3) interacts the commuting indicator with the father's income quartile.

# A4 Return to experience in firm classes

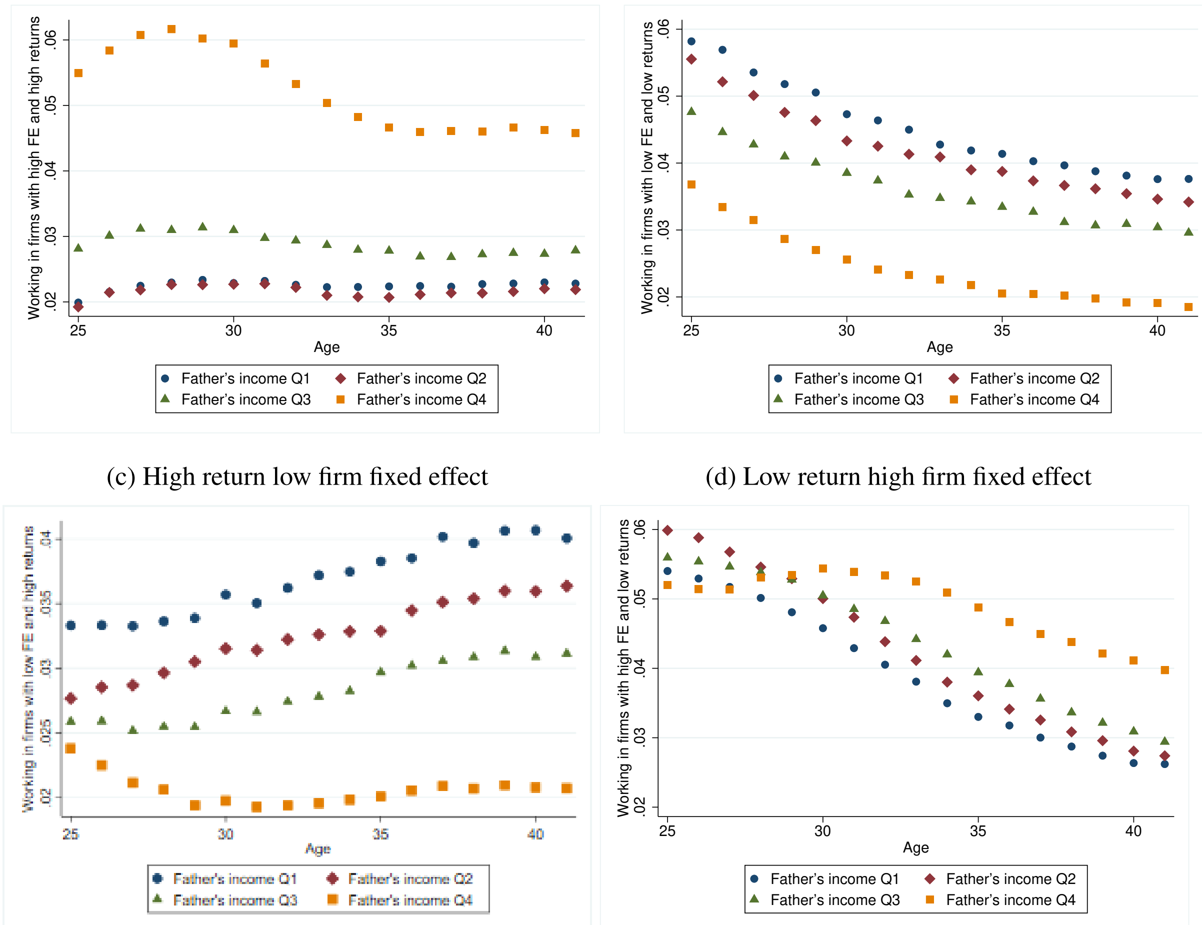
Figure A16: Firm fixed effects and return to experience

(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). Subfigure (a) shows the relationship between firm fixed effects and returns to experience in the firm class. Subfigure (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.

Figure A17: Work in high- vs- low-return firms




(a) High return high firm fixed effect

(b) Low return low firm fixed effects

(c) High return low firm fixed effect

(d) Low return high firm fixed effect

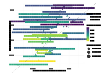
Notes: The figure shows different combination of working in firms with high/low return and high/low firm fixed effects. High (low) return are defined as firms in the 2 groups with highest (lowest) return, and high (low) firm fixed effect are defined as a firm fix effect in the top (bottom) 20 percent.

# A5 Firm values and compensating differentials

The firm values are estimated using the method and code from Sorkin, 2018. Below is a short summary for how the firm values are estimated, for a more detailed description of the estimation see Sorkin, 2018.

In short the firm values are estimated using revealed preferences, using voluntary workers flows between firms. We define voluntary employer-to-employer transitions as moves without unemployment benefits or zero-earnings spells between adjacent firm spells.The value of the firm $Vj = \omega(\psi j + lnaj)$ consist both of a pay component $(\psi j)$ and a non-pay component (aj), and where ω is the (unknown) unit conversion from money to utility.

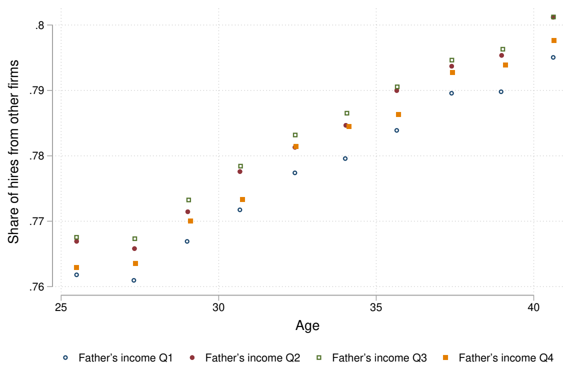
## Figure A18: Industry pay and firm value



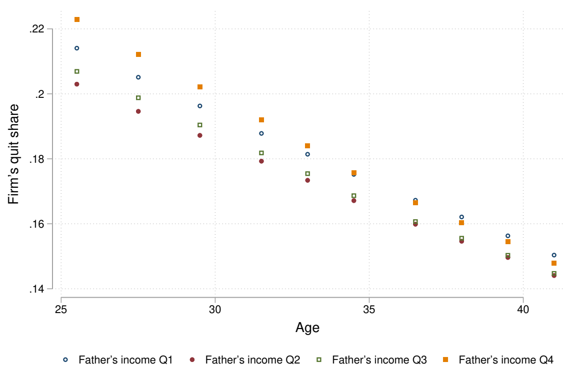
Notes. Mean firm pay and mean firm value in different industries, as calculated from the estimated AKM firm fixed effects based on equation (1) and the mean firm values $\hat{V}_j$ based on Sorkin (2018). The color of the graph shows the share of children from high socioeconomic background in the industry, where high socioeconomic background is defined as those with father income above the median.

## Figure A19: Poaching and quit rates

(a) Poaching rate, controlling for firm pay




(b) Quit rate




Notes: Subfigure (a) shows the mean share of hires from employment (poaching rate), controlling for the estimated firm premium $\hat{\psi}_j$. Subfigure (b) shows the quit rate at the firmm.

Table A12: SES gaps in firm values excluding small firms

| | Dependent variable | | | | | |
|---|---|---|---|---|---|---|
| | $\hat{\psi}_{j=J(i,t)}$ | $\hat{V}_j$ | $\hat{V}_j$ | $\hat{V}_j$ | $\hat{V}_j$ | $\hat{\psi}_{j=J(i,t)}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $y_{f(i)}$ | 0.057*** | 0.045*** | 0.030*** | 0.016*** | 0.011*** | 0.032*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| $\hat{\psi}_{j=J(i,t)}$ | | | | 0.523*** | 0.498*** | |
| | | | | (0.001) | (0.001) | |
| $\hat{\alpha}_i$ | | | 0.128*** | | 0.051* | |
| | | | (0.001) | | (0.001) | |
| $\hat{V}_j$ | | | | | | 0.533*** |
| | | | | | | (0.001) |
| R-squared | 0.029 | 0.019 | 0.072 | 0.293 | 0.300 | 0.300 |
| Observations | 600,108 | 600,108 | 600,108 | 600,108 | 600,108 | 600,108 |

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. Estimated firm values $\hat{V}_j$ are standardized to mean zero and standard deviation one. Small firms are excluded in the table, where small firms are defined as the firms in the two bottoms quantities of numbers of employees. Robust standard errors in parentheses.

Table A13: Correlations between different measures of firms values

| | $\hat{V}_j$ | $\hat{\psi}_j$ | Poaching rates | Quit rates |
|---|---|---|---|---|
| $\hat{V}_j$ | 1.000 | | | |
| $\hat{\psi}_j$ | 0.523 | 1.000 | | |
| Poaching rates | 0.437 | 0.399 | 1.000 | |
| Quit rates | -0.605 | -0.160 | -0.033 | 1.000 |

Notes: Correlation between estimated firm values $\hat{V}_j$ (estimated by following Sorkin, 2018), firm fixed effects $\hat{\psi}_j$ (estimated from equation (1)), poaching rates (defined as the share of hires from employment), and quit rate at the firms.

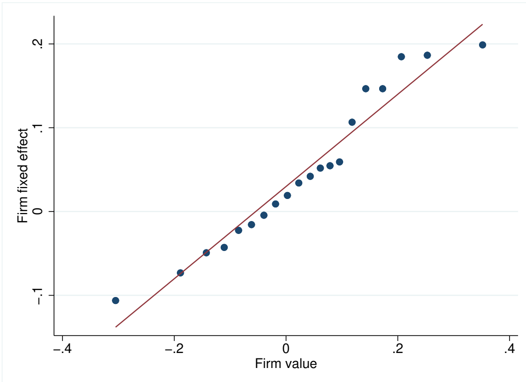
Table A14: Correlations between firm value and firm fixed effect by father's income quartile

| | By father's income | | | |
|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 |
| $Corr(\hat{\psi}_j, \hat{V}_j)$ | 0.509 | 0.497 | 0.513 | 0.526 |

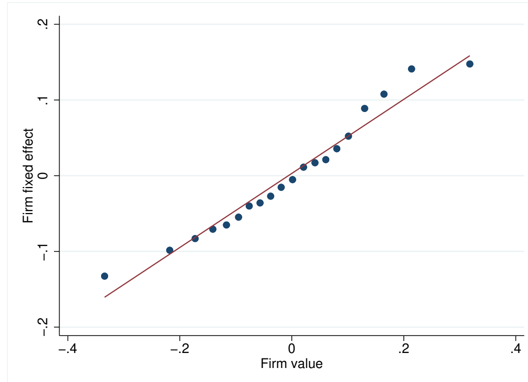
Notes: Correlation between firm fixed effects (estimated from equation (1)) and firm values (estimated by following Sorkin, 2018), by father's income quartile.

Figure A20: Firm value and firm fixed effects

(a) Father's income above the median




(b) Father's income below the median




Notes: The figures plots firm fixed effects (estimated from equation (1)) against binned firm values (estimated by following Sorkin, 2018). Figure (a) includes individuals whose father's income is above the median while Figure (b) includes those whose father's income falls below the median.

# A6 Theoretical derivations

## A6.1 The effect of measurement error on the conditional firm value gradient

Consider the model in Section 5.2 and assume that non-pay amenities $a_j$ are uncorrelated with parental income or firm pay. According to the Frisch-Waugh-Lovell theorem, the coefficient on parental income $y_f$ in a regression of the estimated firm value $\hat{V}_j$ on estimated pay premia $\hat{\psi}_j$ and parental income $y_f$ is equivalent to the coefficient in a regression of $\tilde{V}_j$ on $\tilde{y}_f$, where $\tilde{V}_j$ is the residual from a regression of $\hat{V}_j$ on $\hat{\psi}_j$ and $\tilde{y}_f$ is the residual from a regression of $y_f$

on $\hat{\psi}_j$. We therefore have

$$
\begin{aligned}
\beta_y =&\frac{Cov(\tilde{V}_j, \tilde{y}_f)}{Var(\tilde{y}_f)} = \frac{Cov(V_j - \frac{Cov(V_j,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j, y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)}{Var(\tilde{y}_f)} \\
=&\frac{Cov(\psi_j - \frac{Cov(\psi_j,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j, y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)}{Var(y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)} \\
=&\frac{Cov(\gamma y_f - \frac{Cov(\gamma y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j + \varepsilon_j - \frac{Cov(\varepsilon_j,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j, y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)}{Var(y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)} \\
=&\gamma + \frac{Cov(\varepsilon_j - \frac{Cov(\varepsilon_j,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j, y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)}{Var(y_f - \frac{Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}\hat{\psi}_j)} \\
=&\gamma - \frac{\frac{Cov(\varepsilon_j,\hat{\psi}_j)Cov(\hat{\psi}_j,y_f)}{Var(\hat{\psi}_j)}}{Var(y_f) - \frac{Cov(y_f,\hat{\psi}_j)Cov(y_f,\hat{\psi}_j)}{Var(\hat{\psi}_j)}} \\
=&\gamma - \frac{\gamma Var(\varepsilon_j)}{Var(\varepsilon_j) + Var(u_j^{\psi})} = \gamma\frac{Var(u_j^{\psi})}{Var(\varepsilon_j) + Var(u_j^{\psi})}
\end{aligned}
$$

which corresponds to the expression in the main text.