WORKING PAPER 2026:21

Labor-market drivers of intergenerational earnings persistence

Erika Forsberg
Martin Nybom
Jan Stuhler

IFAU
INSTITUTE FOR EVALUATION OF LABOUR MARKET AND EDUCATION POLICY

The Institute for Evaluation of Labour Market and Education Policy (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala.

IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market and educational policies, studies of the functioning of the labour market and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

More information about IFAU and the institute's publications can be found on the website www.ifau.se

ISSN 1651-1166

Labor-market drivers of intergenerational earnings persistence

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

October 2, 2026

Abstract

How does worker-firm sorting contribute to intergenerational persistence? Firm pay premia account for 36% of the intergenerational earnings elasticity in Sweden, rising to 50% with dynamic returns to firm-specific experience. Large pay gaps already open at labor market entry, as children from privileged backgrounds find more favorable entry points; they then climb the firm pay ladder faster, switching firms more often and gaining more per switch. Skill sorting explains this widening but not the initial gaps, consistent with networks and employer learning. Privileged children's firms are also more desirable overall, a gap that keeps widening after pay gaps stabilize.

JEL Codes: J62, J31, D31

Keywords: Intergenerational mobility, firm pay, worker sorting

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

[†]IFAU (martin.nybom@ifau.uu.se)

[‡]Universidad Carlos III de Madrid (jstuhler@eco.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 Jonas Cederlöf, Elias Einiö, Shahe Emran, Per Engzell, Olle Folke, Rasmus Landersø, Isabel Martínez, Guido Neidhöfer, Forhad Shilpi, Isaac Sorkin, and seminar participants at IFAU, Uppsala University, INSPER, the Tinbergen Institute, Universidad Autonoma de Barcelona, University of Konstanz, University of Innsbruck, the Research Network for Intergenerational Mobility, Workshop on Intergenerational Transmission (Lisbon), University of Oxford, EALE, Verein für Socialpolitik, University of Tampere, and University of Bergen for helpful comments.

Introduction

Children from high-income families go on to 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. The dominant answer points to differences in human capital, with education viewed as a key mediator (Goldthorpe, 2014); seminal contributions in economics emphasize parental investments in human capital (Becker and Tomes, 1979), dynamic complementarities in skill development (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 persistence into pre- and post-birth factors to disentangle nature and nurture (e.g., Björklund et al., 2006; Black et al., 2005). Much of the literature thus 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 persistence.[1]

In this paper, we study how and why worker-firm sorting varies by parental socioeconomic status (SES) and over the lifecycle, and how this sorting contributes to intergenerational earnings persistence. We show that SES differences in firm *pay premia* (henceforth, firm pay gradients/gaps) account for 36% of the intergenerational elasticity of earnings in Sweden, rising to one half once dynamic returns to firm-specific experience are included.[2] Two thirds of the static firm pay gaps emerge at career start, suggesting that high-SES children find more favorable entry points to the labor market. These gaps widen further in early career, as high-SES children climb the firm pay ladder faster, switching firms more frequently and obtaining larger pay gains conditional on switching.

Skill sorting – the tendency of more productive workers to sort into higher-paying firms – explains around half of the firm pay gradient at mid-age and most of its widening over the career, but not the large gaps that emerge already at career start. These patterns may reflect parental and peer networks, but also information frictions: when skills are imperfectly observed, employers initially rely on family background as a proxy for skill, and sorting shifts toward actual skill as information accumulates. Finally, we show that the firm pay gradient reflects a genuine gradient in the overall desirability of jobs rather than compensating differentials: children from high-income families work in firms that are better along both pay and non-pay dimensions, and their advantage in overall firm value keeps growing at ages where the pay gradient has already plateaued.

We start by presenting a set of empirical facts on the socioeconomic gradient in firm

[1] Engzell and Wilmers (2025) make a similar observation from a sociological perspective, arguing that stratification research has largely ignored the role of firms and employers.

[2] Throughout the paper, we use *firm pay gradient* (or *SES gradient in firm pay*) as shorthand for the regression relationship between firm pay premia and parental log income, and *firm pay gap* for mean differences in firm pay across parental-income groups (typically quartiles).

pay in Sweden, following a similar approach as a recent study by Dobbin and Zohar (2025). 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).

We next study how firm pay differences evolve over the lifecycle. Two thirds of the firm pay gap emerges already at career start, consistent with the notion that children from richer families find more favorable "entry points" to the labor market. Their firm pay advantage then increases further in early career before stabilizing in the mid-30s. Differences in early-career progression thus account for the remaining third, driven primarily by differences in the gains conditional on switching job, but also by differences in switching frequency. Viewed through models in which search frictions generate variation in firm pay (Burdett and Mortensen, 1998; Manning, 2003), these findings suggest that high-SES workers climb the "job ladder" more quickly. All these patterns hold within education groups but are especially strong among college graduates.

Moreover, firms shape intergenerational persistence not only through the pay premia they offer today, but also dynamically – through the career progression they enable – a margin that has not been quantified before. Children from advantaged backgrounds sort into firms that grow faster and employ coworkers with steeper earnings growth, pointing to higher returns to accumulated experience. To quantify these dynamic advantages, we follow Arellano-Bover and Saltiel (2026) and Battiston et al. (2024) and estimate an extended two-way fixed effects model in which the returns to experience vary across firm classes. Top-quartile children are about twice as likely as children from the next income quartile to begin their careers in firms that combine high pay premia with high returns to experience. By age 40, the accumulated returns to firm-specific experience add about one fifth to their static firm pay advantage, raising the combined (static and dynamic) contribution of firms to 50% of the IGE. While this estimate should perhaps be interpreted as an upper bound, it implies that the standard "static" perspective substantially understates the role of firms in intergenerational persistence.

We then investigate how much of the SES gradient in firm pay is skill sorting in disguise. Worker and firm effects from the AKM framework correlate positively, as more productive workers sort into higher-paying firms (Card et al., 2013), and children from high-income fam-

ilies have better skills on average. The firm pay gradient may therefore partly reflect skill-based assortative matching, with firms amplifying inequalities that arose already in childhood. Conditioning on the worker effects from the AKM model reduces the SES gradient in firm pay by about 30%. Because estimated worker effects are imperfect proxies of skill, we additionally condition on education and late-adolescence cognitive and non-cognitive skill measures from mandatory (for men) military enlistment tests: observed skills then account for about half of the firm pay gradient at age 40, with the cognitive measure providing the most additional explanatory power. We thus find that skill sorting is a key driver of socioeconomic gaps in firm pay, in line with evidence by Dobbin and Zohar (2025) for Israel and Laliberté and Whalley (2026) for Canada. Our estimates further show that skill sorting is more extensive than what the estimated AKM worker effects alone would capture.

We then trace how the relative importance of skill sorting and family background shifts over the lifecycle. As is intuitive, family background matters most 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%. Skill sorting then gains importance with age, accounting for around half of the firm pay gradient at age 40. Because different mechanisms imply different lifecycle patterns, these dynamics help to discriminate between them. The large role of family background at labor market entry is consistent with a growing literature showing that parents (e.g., Kramarz and Skans, 2014; Staiger, 2026) and peers (Zimmerman, 2019; Cattan et al., 2025) facilitate access to higher-paying firms. But the rising contribution of skill sorting implies that (potentially nepotistic) network effects are not the sole mechanism by which family background affects firm pay.

To interpret these patterns, we formulate a simple model of firm sorting under imperfect information about worker skill, motivated by the literature on employer learning (Farber and Gibbons, 1996; Altonji and Pierret, 2001). If employers use family background as a proxy for skills they observe only imperfectly, statistical discrimination alone can generate our key findings: a large firm pay gradient already at career start, a direct role for family background net of skills, a rising contribution of skill sorting over the lifecycle, and a steepening of the overall gradient with age. Networks may be a related rather than competing mechanism: they may serve as a device to overcome informational frictions (Montgomery, 1992; Hensvik and Skans, 2016), and may thus be most valuable at career start, when such frictions loom largest (Kramarz and Skans, 2014; Willis, 2024).

Finally, we ask whether the SES gradient in firm pay reflects genuine differences in the overall value of jobs, or merely compensating differentials for non-pay job attributes. The distinction matters beyond measurement: if intergenerational transmission partly stems from inherited preferences – for example, if some families place greater weight on income relative to non-monetary attributes – then measured income persistence could over- or understate the persistence of economic welfare. While welfare is not directly observed, revealed-preference

methods allow us to establish the direction of the answer. To our knowledge, there exists little systematic evidence on these questions in the intergenerational context.[3]

Using revealed-preference measures of firm values based on job-to-job flows (Sorkin, 2018), we show that firm pay premia are informative about overall firm desirability – high-paying firms poach more workers and firm values correlate strongly with pay – and that the SES gradient in firm pay extends to a corresponding gradient in firm value: children from high-income families sort into firms that are better not only in pay but also in welfare terms. At a given level of overall desirability, their firms are tilted toward pay rather than amenities – whether reflecting preferences or opportunities – while at a given level of pay, their firms offer more overall value, so their pay premia are less often offset by disamenities. Because firm values have no cardinal scale, the overall SES gradient in amenities cannot be signed; what the evidence establishes is that the firm pay gradient reflects a genuine gradient in firm desirability, rather than compensation for disamenities. Moreover, the SES gradient in firm value keeps widening at ages where the pay gradient has already plateaued: the pay–amenity trade-off shifts with age, and more so for high-SES children.

Related literature. Our work contributes to a nascent literature on the contribution of labor market sorting to intergenerational persistence. A key motivation is Dobbin and Zohar (2025), whose methodological approach we adopt in our baseline analysis. Closely related are recent or concurrent papers by Engzell and Wilmers (2025), who take a sociological perspective grounded in stratification theory to study the transmission of firm pay advantages in Sweden, Laliberté and Whalley (2026), who analyze the role of social connections in firm pay advantages in Canada, and Card et al. (2026), who compare the roles of parents and neighborhoods for the transmission of firm pay in the US. 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 strata. Laliberté and Whalley distinguish the role of connections from skills by conditioning on fine-grained educational information, concluding that social connections account for about 15% of the SES gradient in firm pay.

We extend this literature by tracing the evolution of pay gaps over the lifecycle, by capturing the dynamic implications of worker-firm sorting, by using direct skill measures to study skill sorting over the lifecycle, and by examining the role of non-wage attributes in overall firm desirability. In comparison, we find a larger contribution of firms to intergenerational persistence than other studies, explaining 36% of the IGE at mid-age in the static model and about *half* in the dynamic model. This reflects not just cross-country variation, but also differences in data quality, our use of establishment identifiers for the largest firms, and – in the

[3] 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.

dynamic model – the incorporation of firm-specific returns to experience.

Our work also relates to a growing literature on social connections in worker-firm sorting. In a pioneering study, **?** 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 (2026) 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 strand highlights the importance of peer networks formed in educational or other settings.[4] Others emphasize broader concepts of social capital, including as a sources of socioeconomic disparities in labor-market outcomes (Putnam, 1995; Chetty et al., 2022). Particularly relevant for our lifecycle analysis, Stansbury and Rodriguez (2026) show that differences in social capital contribute to SES gaps in career progression within U.S. academia.

The "networks" and the "sorting" literatures are complementary. The former identifies the causal effect of specific connections, but cannot quantify their overall contribution, because network composition and strength are not fully observed. The latter quantifies the total contribution of labor market sorting to intergenerational persistence, but cannot separate the role of networks from other mechanisms. Our lifecycle perspective and direct skill measures allow us to bring the two closer. Our finding that large firm pay gaps emerge already at labor market entry – and that these gaps are not accounted for by individual skill – is consistent with networks shaping entry points into the labor market. The subsequent widening of the gradient, in contrast, is better explained by employer learning under imperfect information about worker skill. As we show, such information frictions would also amplify the effect of family background on initial pay gaps, perhaps helping to explain their large magnitude.

Finally, our work relates to recent studies of skill sorting in the labor market (Eeckhout, 2018; Card et al., 2018). Consistent with Dobbin and Zohar (2025), we find far more sorting by parental background than skill sorting alone would predict. We further document firm sorting by cognitive and non-cognitive skills (similar to Nybom, 2017, for college education), over and above the sorting captured by the AKM worker effects. Differences in firm premia also 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 accounts for 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 in SES pay gradients in Sweden. A recent literature highlights the contribution of labor market

[4] Zimmerman (2019), Michelman et al. (2022), and Barrios-Fernández et al. (2024) show that admission to elite colleges improves labor market outcomes primarily through changed social environments rather than academic achievement. Campa (2025), Cattan et al. (2025), and Fischer et al. (2025) find that exposure to more affluent schoolmates improves long-term outcomes, and Einiö (2026) that quasi-random exposure to high-income peers among conscripts raises long-term earnings.

sorting to cross-sectional inequality (7–12% in our data, comparable to estimates for other countries in Bonhomme et al., 2023). Our findings suggest, however, that firm pay differences account for a considerably larger 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.[5] Firm and establishment characteristics, including location and industry codes, are available from 1985 onward.

Main intergenerational sample. Our primary sample includes men and women born between 1967 and 1977, followed on the labor market from ages 25 to 41, for whom we observe at least one identified firm effect. For these cohorts we can observe a substantial portion of their careers while also measuring their fathers' income in prime age. Some analyses, however, focus on peak-career outcomes at ages 39–41. Fathers' long-run income is measured between ages 45 and 55 using either mean log earnings or mean earnings rank; 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.

AKM sample. To estimate firm premia, we use a matched employer–employee dataset covering the full Swedish labor force ages 20–64 from 1985 to 2018. Firms are the unit of analysis, except for multi-establishment firms with more than 1,000 employees (on average over the period), for which establishment identifiers are used.[6]

Descriptive statistics. Our final samples comprise 836,743 individuals observed at ages 39–41, and 960,925 individuals observed at some point between ages 25 and 41. Appendix Table A1 reports descriptive statistics for both samples and details the construction of the

[5] 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.

[6] We also show alternative versions of our baseline results using only firm or only establishment identifiers.

sample; differences relative to the AKM sample mainly reflect age and cohort composition.

1.2 Estimation of worker and firm 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 α_i denotes an individual (or "worker") fixed effect, ψ_j a firm fixed effect (the "firm pay premium"), $\boldsymbol{X}_{it}$ a vector of time-varying controls with coefficient vector $\boldsymbol{\delta}$, and ε_{ijt} an error term. The controls in $\boldsymbol{X}_{it}$ include year dummies and education-gender-specific age dummies.[7] 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 restrict age effects to be constant over ages 45–54, where earnings profiles are roughly flat.

Estimating the firm pay premia ψ_j by OLS corresponds to pairing eq. (1) with a strict exogeneity restriction on the errors ε_{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.[8]

We estimate equation (1) using our AKM sample for the years 1985–2018, focusing on workers with substantial labor-market attachment as approximated by the earnings threshold described in Section 1.1. As the firm pay premia are identified from earnings changes when workers switch firms ("movers"), they are comparable only within a set of firms connected through such transitions. We therefore compute the largest connected set for our time period and drop worker-year observations associated with firms outside this set (about 0.1% of observations).

Firm effects are imprecisely estimated for firms connected by only a few movers (Andrews et al., 2008; Kline et al., 2020; Bonhomme et al., 2023), which inflates their variance and biases their covariance with the worker effects ("limited-mobility bias"). Classical mea-

[7] 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.

[8] These assumptions are unlikely to be fully met in practice. Nevertheless, the linear specification in eq. (1) can serve as a useful approximation, 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 account for only a small share of the variance of log earnings (Bonhomme et al., 2019; Kline, 2024). Most relevant for our setting, Engzell and Wilmers (2025) show that in Sweden, the earnings gain from working at a high-paying firm is similar for workers from high and low-income family backgrounds.

surement error would not affect our main analysis, as the estimated firm effects serve as the dependent variable; however, as some analyses use them as a covariate, we drop firms connected by fewer than five movers, removing an additional 6% of worker-year observations.[9]

In the subsequent sections, we use the estimates of α_i and ψ_j as inputs in our intergenerational analysis. The worker effects represent the portable component of earnings, reflecting a worker's productivity but possibly also other permanent factors, such as bargaining power or discrimination at the market level (Kline, 2024). The firm effects represent a non-portable component of earnings enjoyed only when employed at firm j. Such firm pay premia may arise from search frictions (Burdett and Mortensen, 1998), but may also reflect match-specific productivity, amenities ("compensating differentials"), or external constraints such as collective bargaining agreements (Gibbons et al., 2005; Sorkin, 2018). Intergenerational earnings persistence may operate through both the worker- and firm-specific components.

Firm pay premia vary systematically with firm characteristics. Average premia are highest in oil and natural gas extraction, financial services, telecommunications manufacturing, chemical manufacturing, and IT services, and lowest in hotels and restaurants, domestic work, agriculture, forestry, and retail trade. Consistent with prior research, firm pay is positively correlated with firm size ($r = .15$), with the pay–size gradient particularly steep at the lower end of the size distribution. Average firm pay is also substantially higher in the private sector than in the public sector (means of .06 versus –.06). Geographically, firm pay tends to be highest in Stockholm and the surrounding counties, as well as along the West Coast – and lowest in more rural counties.

Appendix A1.2 reports the variance decomposition of log earnings implied by equation (1) for both the AKM and the intergenerational samples. Worker effects account for 30–38% of the variance, firm effects for 8–13%, and their covariance for another 7–12%; the implied correlation between worker and firm effects of 0.40–0.55 reflects strong assortative matching. These results are similar to Engbom et al. (2023), and broadly consistent with evidence for other countries (e.g., Song et al., 2019; Bonhomme et al., 2023).

2 Firm pay premia and 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

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

[9]Table A6 shows that the firm pay gradient increases slightly when we include firms with fewer than 5 movers, reflecting differential sorting into small, low-paying firms.

where y_{ijt} is child log earnings in year t, $y_{f(i)}$ the log earnings of the father of child i, and β 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 α_i. Conversely, preferential labor-market access through parental networks may manifest as a positive association between parental earnings and the firm effects ψ_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 β. Our primary focus is on the slope coefficient $\beta_{\textit{firm}}$ from regressing the child's estimated firm pay premium $\hat{\psi}_j$ on parental log earnings $y_{f(i)}$,

$$\hat{\psi}_j = \mu_\psi + \beta_{\textit{firm}} y_{f(i)} + \omega_{ijt}, \tag{3}$$

which we refer to as the SES gradient in firm pay (or simply, the *firm pay gradient*).

Table 1 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 for *labor* income in pooled samples of 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 (or "worker") effects, which capture all permanent worker heterogeneity, including but not limited to productive 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.

Comparison to previous studies. The contribution of worker-firm sorting to *cross-sectional* inequality has received much attention (Card et al., 2013; Bonhomme et al., 2023), but sorting accounts for a much larger share of *intergenerational* inequality. In our data, sorting explains about 7–12% of the variance in log earnings, depending on sample (Appendix A1.2)[13] – roughly a third of the worker component – but 36% of the IGE, more than half of the worker component. Our estimate exceeds the 23% that Dobbin and Zohar (2025) attribute to firms

[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]Although second-step regressions of estimated firm effects on worker covariates yield unbiased coefficients, inference is complicated by correlation between the fixed-effects estimates (Kline, 2024). Using two years of income data, Kline (2024) shows that standard errors based on the cross-fitting procedure of Kline et al. (2020) are about 75% larger than naive ones. This is less of a concern here, given the small standard errors in Table 1, and the far larger number of person-year observations underlying our estimated firm effects.

[12]The correlation between the log of father's earnings 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 1: Decomposition of the intergenerational earnings elasticity

	Dependent variable				
	y_{ijt} (1)	$\hat{\alpha}_i$ (2)	$\hat{\psi}_{j=J(i,t)}$ (3)	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$ (4)	$\hat{\psi}_{j=J(i,t)}$ (5)
$y_{f(i)}$	0.198	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 the 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 worker effects α_i, mean firm effects ψ_j for ages 39–41, and time-varying controls. Robust standard errors in parentheses. Column (5) adds the estimated worker effect as a control; its coefficient is reported in the second row.

in Israel, the main estimates of Engzell and Wilmers (2025), who use similar Swedish data but coarser firm identifiers and earnings ranks rather than levels,[14] and the roughly 15% that Card et al. (2026) attribute to establishment effects in the US.

Why do we find stronger worker-firm sorting by family background in Sweden than recent work found for the US and Israel? Sorting may be genuinely stronger in Sweden: low returns to skill and a homogeneous, tuition-free education system may raise the importance of firm sorting *relative* to time-invariant skills. Differences in data quality are another possibility.[15] However, because the estimated firm pay premium is the dependent variable in eq. (3), classical measurement error in the premia does not introduce bias,[16] and measurement error in parental income attenuates the IGE but not the relative contribution of firms.

A key factor, however, is the definition of firms and their coverage, as both firm pay and sorting by family background vary with firm size. Children from low-income families are more likely to work in small firms with particularly low premia, and such firms are the first to drop out of the connected set in less representative samples. A second issue, specific to

[14]When Engzell and Wilmers (2025) use establishment ("workplace") identifiers the firm contribution to the rank correlation rises from 23% (sum of their firm-to-firm and worker-to-firm contributions in their Table 1) to 31% for sons and 35% for daughters (their Table A3). Moreover, their 23% baseline estimate rises to 28–29% when earnings are measured in logs rather than ranks (their Table A6).

[15]The data differ in several respects. First, we measure parental earnings over a comparatively long mid-life age span, age 45–55. Second, our AKM estimation covers a longer period (1985–2018) than the Israeli data (2010–2015). Third, among children born 1967–1977 with observed paternal earnings, our earnings and firm-effect restrictions retain about 90% (Appendix A1.1), while Dobbin and Zohar drop about half of their sampled cohorts due to insufficient information on earnings or its AKM components.

[16]In line with this claim, randomly dropping half of our individuals before estimating the AKM model has little effect on our estimates (Appendix Table A5), if the sample of firms is held fixed.

short AKM windows, is the contamination of annual earnings in "transition years" in which workers switch employers: such earnings reflect pay at two firms, or only part of a year, and in a shorter window they make up a large share of the observations that identify each firm's premium. Consistent with both mechanisms, the firm pay gradient decreases when we raise the minimum number of movers per firm (Appendix Table A6, Panels D–F) or shorten the AKM window; short- and long-window estimates converge once transition years are dropped and the firm sample is held fixed (Appendix A1.3.2). Finally, using establishment identifiers for large firms raises the gradient considerably (Appendix Table A6, Panels B and C).[17] Section 2.3 provides further sensitivity evidence.

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, with worker and firm fixed effects positively correlated (Card et al., 2013; Song et al., 2019). This matching contributes to cross-sectional income inequality (eq. (A1) and Appendix Table A2). It may also contribute to intergenerational persistence, since worker effects are correlated with parental income (Table 1, 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.

As a first illustration of the role of assortative matching, we regress the firm fixed effects on father's log income *conditional* on the worker fixed effects (Table 1, column 5). The coefficient on father's log income declines by nearly 30%, from 0.065 to 0.047, and the coefficient on the worker effect is large and statistically significant – reflecting substantial assortative matching. This is in line with Dobbin and Zohar (2025), who find an even larger decline in the firm pay gradient for Israel. Part of the firm pay gradient thus reflects the SES gradient in permanent worker characteristics: worker-firm sorting amplifies the pay-off to the well-known SES gradient in human capital. Skill sorting is difficult to measure precisely, however, and we return to this question in Section 5.

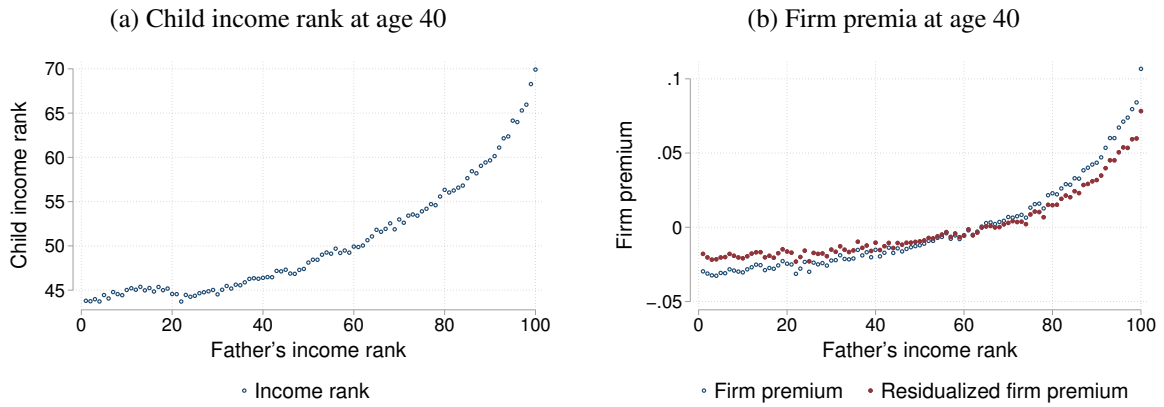
2.1 Non-linearities in the firm pay gradient

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. Both figures reproduce the positive intergenerational relationships documented in Table 1, 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 worker fixed effect (red circles in subfigure b), but smaller in magnitude. Moreover, the gap between the unconditional (blue circles) and conditional (red circles) gradients widens toward the top

[17] Our baseline splits firms with more than 1,000 employees into their establishments (see Section 1.1).

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



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 ψ_j at age 40 as estimated by equation (1) and firm premia residualized on worker fixed effects.

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. While the non-linear pattern is less stark when using logs instead of ranks, the growing importance of sorting at the top holds across specifications (see Appendix Figure A4).

2.2 The role of region, industry, and firm size

Since pay premia vary systematically with firm characteristics (see Section 1.2), the firm pay gradient documented in Table 1 may the reflect the sorting of high-SES children into particular regions and industries or specific firm types. Table 2 explores this possibility by sequentially controlling for different sets of firm characteristics. 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 a similar reduction of 33%. Conditioning on sector (public vs. private) has only a minor impact (column 4), while firm size has a moderate effect on the estimated SES gradient, reducing it by 17% (column 5). Jointly controlling for region, industry, sector, and firm size explains 61% of the unconditional SES gradient in firm pay (column 6), while occupations explain part of the remainder (Table A8). Thus, an important share of the "firm" pay gradient reflects observable firm characteristics.

The margins that contribute most to the SES gradient in firm pay are not necessarily those that contribute most to the overall variance in firm premia. Regions and industries account for similar shares of the SES gradient, but industries explain a much larger share of the variance

Table 2: 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					X	X
Share of β_{firm}	100%	30%	33%	2%	17%	61%
R^2	0.032	0.139	0.389	0.166	0.154	0.524
Worker obs.	783,924	783,924	783,924	783,924	783,924	783,924

Notes: Column (1) reports estimates of the slope coefficient from regressing $\hat{\psi}_j$, at age 40 as estimated from equation (1), on log father's 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 firm size (10 deciles). Robust standard errors in parentheses.

in firm premia (R^2 of 0.389 vs. 0.139). Conversely, firm size and the public sector indicator explain similar shares of the firm pay variance, but only firm size contributes to the SES gradient (17% vs. 2%). The regional and firm size distribution of high-SES children are thus key margins for their pay advantage, while sorting into the public sector plays no major role.

2.3 Sensitivity and heterogeneity

We examine subgroup differences and the sensitivity of our estimates to specification and sampling choices. First, we verify that the estimated role of firms remains similar when parental income is measured as the average of both parents' incomes rather than the father's alone (Table A3).[18] Second, because our long estimation window may strain the assumption of time-invariant firm premia, we allow firm effects to vary across four subperiods of the 1985–2018 period (e.g., Lachowska et al., 2023; Engbom et al., 2023); this modestly increases the estimated role of firms (Table A6, Panel A). Third, we consider sorting across establishments rather than firms. Establishment fixed effects explain 34% of the IGE (37% after netting out AKM residuals, Panel B), slightly more than with our baseline definition that distinguishes establishments only for large firms (see Section 1).[19] Conversely, always using firm identifiers regardless of firm size reduces the estimated firm contribution, mainly driven

[18]The baseline IGE is substantially larger when using the log of average parental income, but this mainly reflects its lower variance; the corresponding correlations are similar (0.23 vs. 0.20).

[19]Many smaller firms consist of only one establishment. Using establishments increases the number of "firm units" from 460,875 to 469,625.

by public-sector "firm units" (Panel C). Fourth, we vary the minimum number of "movers" required for each firm in the AKM estimation. Restricting the sample to firms with at least 10 or 50 movers reduces the firm pay gradient (Panels E and F) compared to retaining all firms (Panel D). Excluding firms with few movers tilts the sample toward larger firms, where a single firm-wide premium is less informative about individual pay; we therefore retain the baseline threshold of five movers.

Across subgroups, the firm pay gradient as a share of the IGE is very similar for women and men (Table A7, Panels A and B). While around 5% of children work in the same firm as their father (in line with **?**; Stinson and Wignall, 2018; Laliberté and Whalley, 2026), dropping these "firm followers" leaves the firm pay gradient virtually unchanged (Panel C). In contrast, excluding public sector workers raises the firm contribution by nearly 10% (Panel D), consistent with more compressed and homogeneous pay-setting in the public sector.

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 worker fixed effects, and that such firm-sorting accounts for one third 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 firm pay gap already at age 25, with children from higher-income families working at higher-premium firms. Moreover, those from the top quartile (yellow squares) experience a notably steeper gain in firm premia early in their careers: while children from the bottom three quartiles see their premium rise by about 1 percentage point by their early 30s, the increase among the top quartile is around 3 percentage points. The gaps stabilize from the mid-30s onward. Given a standard deviation of firm premia of around 0.15 (see Table A2), the advantage of the top quartile over the next corresponds to about 30% of a standard deviation. These lifecycle patterns are qualitatively similar for men and women (Appendix Figure A7), though average firm premia are consistently higher for men.[20]

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-

[20] While the estimated firm fixed effects are mean-zero in the full AKM sample, they tend to be positive in our intergenerational sample.

income families. Figures 2b and 2c therefore reproduce the analysis separately for children with at most a high school degree and those with some college or more. Clear SES gaps in firm pay are evident in both groups. However, among non-college workers, firm premia rise less steeply over age. Among the college-educated, premia rise across all quartiles, but the gains are notably larger for the top quartile. The qualitative patterns are otherwise similar across education groups.[21] We therefore focus on the full sample, reporting education-specific results in the Appendix and discussing them where relevant. The qualitative patterns also hold when we condition on worker fixed effects to account for skill-based assortative matching (Appendix Figure A6, the lifecycle analogue of column 5 in Table 1).

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 standard AKM assumptions, pay otherwise identical workers more. This gap widens during the early career and stabilizes by the mid-30s. In Section 5, we show that the initial gap is largely unexplained by differences in worker skills, while its subsequent widening almost entirely reflects skill differences that correlate with parental background.

3.2 Climbing the firm ladder

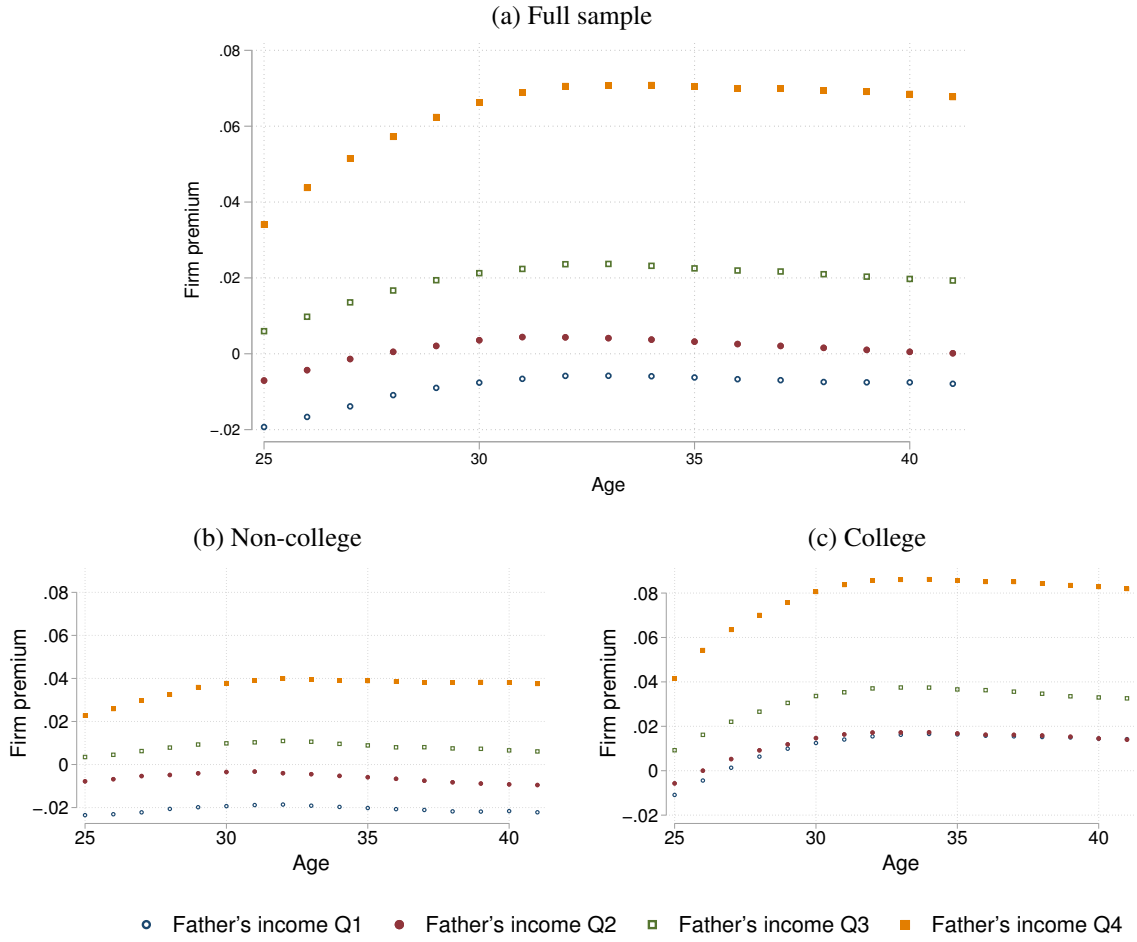
While high-SES children tend to start their careers at higher-paying firms, the firm pay gaps also magnify considerably during their first 10–15 years in the labor market. This widening may reflect either more frequent firm switches among high-SES individuals (climbing the job ladder faster), more advantageous switches (the rungs of their ladder are further apart), or both.

Figure 3a shows the annual firm-switching rate, i.e. the probability of being observed in a new firm at age t relative to age $t-1$, 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 pattern holds within education groups (Appendix Figure A9), with higher mobility overall among the college educated and more pronounced SES gaps 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

[21] This similarity reflects that time-constant differences in worker pay are captured by the worker 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 Figure A5 in Appendix A2).

Figure 2: Firm pay premium over the lifecycle



Notes: Mean estimated firm premium $\hat{\psi}_j$ over the lifecycle, 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.

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 tends to even out, and the SES gaps diminish. 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 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.[22]

[22]The larger gains from firm switching among high-SES children cannot be explained by unemployment dynamics. To show this, we use UI benefit receipt to distinguish voluntary switches (no UI benefit receipt between employment spells) from involuntary switches (those that occur through a spell of unemployment). As shown in Figure A12, high-SES children are more likely to experience an improvement in the firm premium

This difference narrows with age, partly for mechanical reasons: workers who have already reached high-paying firms have less scope for further gains ("ceiling effects"), so that under random switching, low-SES workers would experience larger average gains from their lower starting points. Figure 3d therefore conditions on the firm effect in the previous year; with this adjustment, high-SES children continue to gain systematically more from firm switching throughout the lifecycle.

Appendix Figures A10 and A11 replicate these analyses separately by education group. College-educated individuals are generally more likely to experience premium-improving switches and larger gains in firm pay throughout their careers, yet the SES gaps are qualitatively similar across education groups.

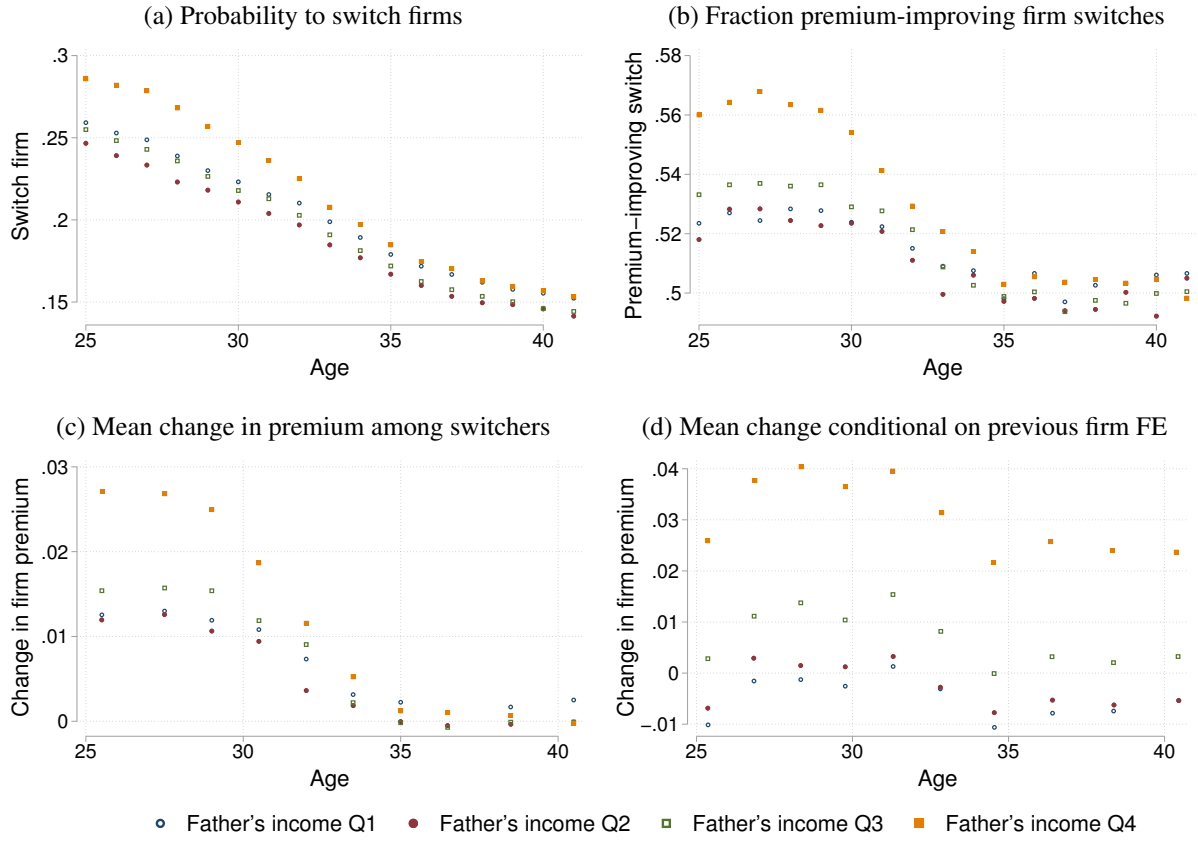
To quantify the relative contributions of switching frequency and switch quality, we decompose the widening of the Q4–Q1 firm pay gap between ages 25 and 40. A Shapley decomposition shows that differences in switch quality account for roughly 90 percent of the widening in the firm pay gap, whereas differences in switching frequency account for the remainder.[23] Thus, the lifecycle divergence in firm pay primarily reflects more advantageous moves among high-SES workers, rather than greater mobility per se.

Finally, we consider the role of commuting. Children from high-income families may have stronger networks and more flexible transport options, expanding the set of feasible jobs. Indeed, the commuting rate is about 8 percentage points higher in the top than in the bottom parental-income quartile, and commuters tend to work in better-paying firms. In event-study estimates, however, starting to commute raises firm pay by only 1–2 percentage points, with similar gains across parental backgrounds. Differences in commuting thus contribute little to the firm pay gaps (see Appendix A2.2).

following both voluntary and involuntary switches.

[23] We construct counterfactual Q4 trajectories by replacing, in turn, the Q4 switching rate and the Q4 gain per switch with their Q1 values. Because the two enter multiplicatively, the split depends on the replacement order (88.1% vs. 93.5% for switch quality), so we report the Shapley average over both orders: 90.8% for switch quality and 9.2% for switching frequency.

Figure 3: Firm switching patterns over the lifecycle



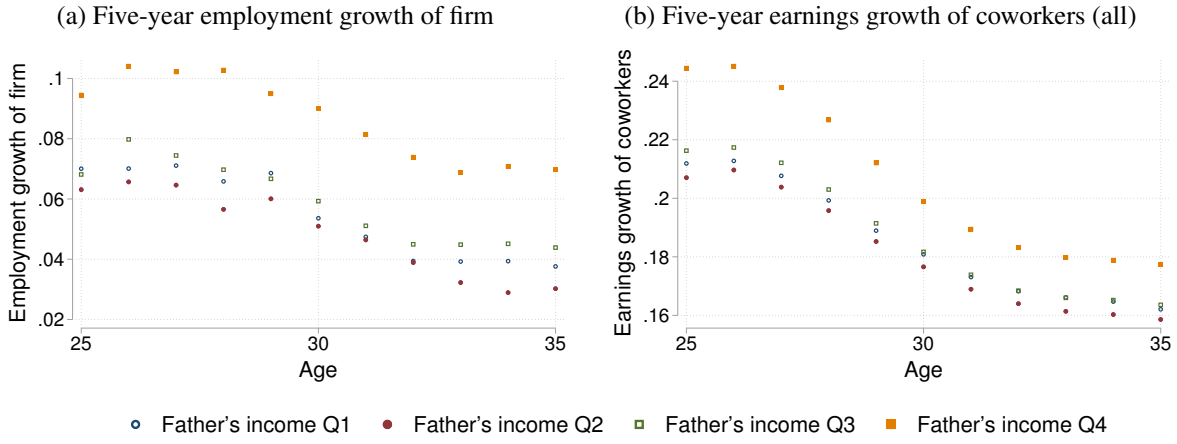
Notes: Patterns of firm switching over the lifecycle, by quartile of father's income. Subfigure (a) shows the probability of working in a different firm than in the year before. Subfigure (b) shows, conditional on switching, the probability of moving to a firm with a higher firm premium. Subfigure (c) 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.

4 Beyond "static" firm pay premia

So far we have characterized firms by their "static" pay premia. Yet firms differ along many other dimensions that shape careers: the productivity of their workforce, how fast they grow, and the returns to experience they offer. This section shows that children from high-income families also sort into firms that are advantageous in these "dynamic" respects – and that accounting for dynamic advantages raises the contribution of firms to intergenerational persistence further.

To illustrate such dynamic differences, Figure 4 shows that children from high-income families work in firms that grow faster – over the subsequent five years, the firms of top-quartile children add employment 40–100% faster than those of children from the lower quartiles (subfigure a) – and whose workers experience steeper earnings growth (subfigure b). This difference partly reflects coworkers' mobility to other firms rather than earnings

Figure 4: Dynamic firm characteristics



Notes: Mean difference in firm characteristics between year t and year $t+5$ for the individuals who work in the firm at each indicated age, by quartile of father's income. Subfigure (a) plots the five-year employment growth of the firm; subfigure (b) plots the five-year earnings growth of all coworkers, including those who stay and those who switch firm.

growth within firms (Appendix Figure A8). High-SES children also have more productive and more affluent coworkers – mean AKM worker effects are about 6 percentage points higher in the top than in the bottom parental-income quartile, and the share of high-SES coworkers is higher (Appendix Figure A8) – a segregation that is most pronounced in early career and that may shape trajectories through learning spillovers, bargaining power, or social networks (Kremer, 1993).[24]

4.1 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 coworkers 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 coworker earnings growth may reflect better learning or promotion opportunities. Even firms that do not pay particularly well today could be attractive if they offer high returns to accumulated experience, raising earnings later in the career.

To examine the role of such firm-specific returns to experience, we follow Arellano-Bover and Saltiel (2026) and Battiston et al. (2024). We begin by randomly splitting our sample into two groups. In one subsample, we classify firms into ten classes based on the distribution of

[24]Interestingly, gaps between the top parental income quartile and the rest in coworker earnings growth and firm employment are pronounced and relatively stable over age, while differences among the bottom three quartiles are generally much smaller.

stayers' annual unexplained earnings growth, using a k-means clustering algorithm.[25] We then use the second subsample to estimate firm-class-specific returns to experience. Specifically, we estimate an extended two-way fixed effects model:

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

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, so that γ_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.[26]

Figure 5 summarizes the results. Panel (a) shows the estimated return to an additional year of experience in each firm class m, where class 1 contains firms with the highest returns and class 10 the lowest. As in Arellano-Bover and Saltiel (2026) and Battiston et al. (2024), we find substantial heterogeneity across firm types, though the differences in firm-specific returns are smaller in our Swedish data than in Italy or (especially) Brazil: workers in the top firms (class 1) enjoy returns of about 2 percentage points per year (net of education-by-gender specific earnings growth captured by $\boldsymbol{X}_{it}\boldsymbol{\beta}$), while those in class 10 see slightly negative returns.[27]

Figures 5b and 5c plot the probability of working in high-return firms (classes 1 or 2) and low-return firms (classes 9 or 10), by parental-income quartile and age. Children from high-income families are more likely to work in high-return firms early in the career: about 14% of top-quartile children work in high-return firms, versus 10–11% of others. They are also 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.

The lifecycle profile of working in high-return firms stands out. Whereas low-SES children become increasingly likely to sort into these firms over their careers, those from the top

[25] Unexplained earnings growth is stayers' year-on-year earnings growth net of year and education-by-gender-by-age effects. We characterize each firm by the shares of its worker observations below 20 roughly equally spaced percentiles of the pooled distribution – a 20-point approximation to its CDF – and group these with k-means into ten clusters, so that firms are classified by the shape of the whole distribution rather than by its mean. Clustering at the worker-firm observation level gives firms with more observations more weight.

[26] Unlike the static AKM model, we here restrict the sample to workers who we can track to age 41, corresponding to cohorts born between 1967 and 1977 (as in our intergenerational regressions). Since we observe individuals only to age 41, we cannot impose that age effects are constant over 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 age polynomials interacted with education and gender.

[27] Like static premia (Section 1.2), returns tend to be higher at larger firms, and in industries such as oil and gas extraction and financial services. Unlike static premia, the lowest returns are found mainly in industries in manufacturing and returns are, if anything, slightly higher in the public than in the private sector. Geographically, returns are highest in the Stockholm and neighboring Uppsala regions, as are static premia, but the patterns otherwise differ.

quartile do the opposite: while being substantially more likely to work in high-return firms early on, they increasingly sort out of them, and by their late 30s are less likely to work in them than children from the bottom half of the distribution. This finding points to potential trade-offs, with high-SES children beginning in high-return firms that offer strong career prospects and later moving to firms that may offer other, potentially non-monetary benefits. We return to this question in Section 6.

Figure 5d summarizes the overall pattern, plotting the 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 thereafter, as they sort away from the highest-return firms. Up to age 30, their firm-specific returns are on average about 40% higher than those of their lower-SES peers. Interestingly, it is those from the second-lowest rather than the lowest quartile who experience the lowest firm-specific returns to experience.

Returns and pay premia need not go together: for some firm classes there is a trade-off between "dynamic" returns and "static" pay premia, while others combine high returns with high premia, or low returns and low premia.[28] Figure A17 shows how low- and high-SES children distribute across the four combinations of high/low static and dynamic advantages. As expected, high-SES children are more likely to work in firms offering both high premia and high returns, and low-SES children in firms with low returns and low premia. The magnitude of these gaps is striking, with top-quartile children twice as likely as those in the next quartile to work in "high-high" firms early in their career. The off-diagonal patterns are also notable: children from high-income families are less likely to begin their careers in firms with low dynamic returns but high static premia, yet much *more* likely to work in them at later ages.

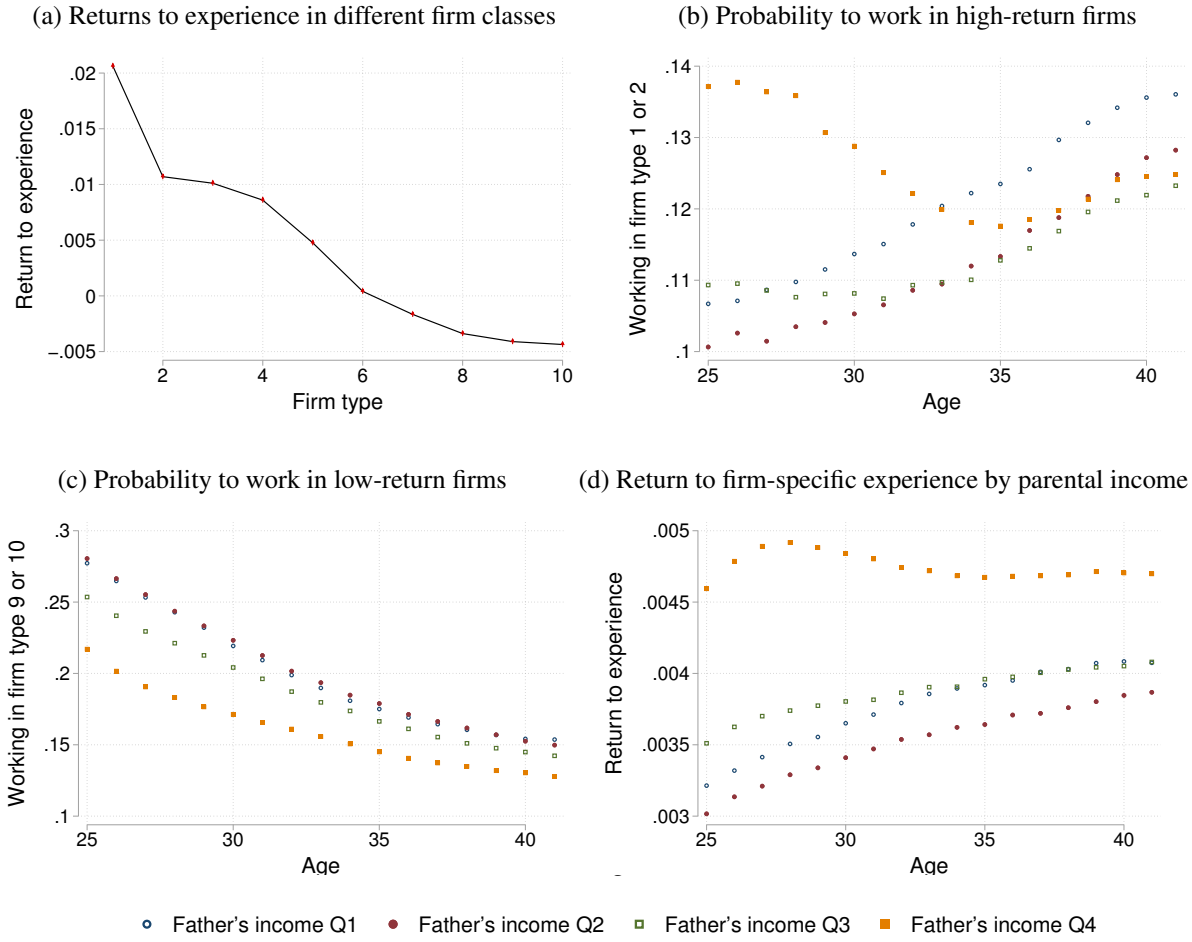
Overall, these patterns show that high-SES children sort into high-growth firms early in their careers, but later become overrepresented in lower-growth – but high-premium – firms. They thus find more favorable entry points to the labor market not only in a static sense but also in a dynamic one.

4.2 The contribution of dynamic firm advantages to the IGE

Table 3 quantifies how much of the intergenerational earnings elasticity at age 40 can be attributed to the static and dynamic firm advantages in equation (4). 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

[28] As shown in Appendix Figure A16a, static pay premia and dynamic firm returns are only weakly correlated across firms: firm classes with similar returns can differ substantially in their pay premia. Arellano-Bover and Saltiel (2026) find the same pattern in data from Rio de Janeiro and Veneto.

Figure 5: High- vs. low-return firms



Notes: Subfigure (a) shows estimates of the coefficients on firm-class experience from equation (4). 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.

in firm pay over age (i.e., climbing the firm ladder, see Section 3.2). Additionally, sorting into firms with higher *returns to experience* explains another 8% of the IGE (column 5). Appendix A3 provides further detail on these firm classes, including a decomposition of the relationship between returns to experience and parental income (Table A10).

Because the return to firm-specific experience is identified by comparing workers at the same firm with different prior experience, it does not capture returns that may materialize only upon moving to other, better-paying firms. In that sense, the contribution in column (5) of Table 3 can be seen as a lower bound on the overall contribution of dynamic firm advantages to the IGE. Indeed, changes in firm premia are positively correlated with the estimated returns γ_m in the firm class an individual switches from (Figure A16b), suggesting that firm-specific experience also helps explain why high-SES children climb the job ladder more quickly. This dynamic perspective thus implies an even larger role for firms than the static decomposition

Table 3: 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 slope coefficient from regressing log child earnings at age 40 on log father's earnings. Columns (2)–(6) report slope coefficients from the corresponding regressions for the components of child earnings in equation (4): the worker fixed effect α_i, the firm fixed effect at age 25, its change between ages 25 and 40, returns to firm-specific experience (the cumulative sum of γ_m over ages 20–40), and time-varying controls. The sample is restricted to workers born 1967–1977 who are observed in a firm at age 25, excluding the random half used to classify firms. Robust standard errors in parentheses.

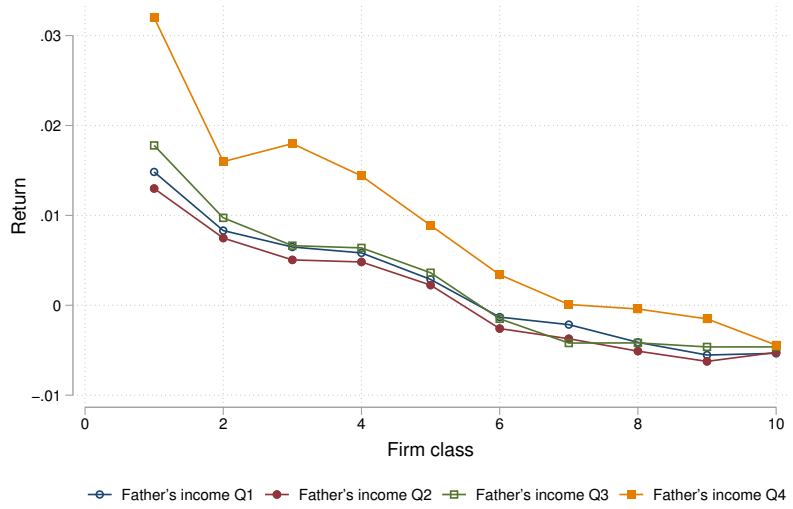
in Table 1: combining columns (3)–(5), firms account for about 46% of the IGE, or 50% when netting out the contribution of other covariates (column 6).

A potential concern is that children from high-income families may experience higher returns to experience regardless of which firms they sort into. The differences in returns across firm classes in Figure 5a would then partly reflect heterogeneity across individuals rather than firms, and the 50% share should then instead be treated as an upper bound. Following Arellano-Bover and Saltiel (2026), who interact firm-specific experience with worker fixed effects, we allow the returns to experience to vary with parental income, estimating

$$y_{ijt} = \alpha_i + \psi_j + \sum_{m=1}^{K} \gamma_m Exp(m)_{it} + \sum_{m=1}^{K} \delta_m Exp(m)_{it} * \theta_i + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijt}, \tag{5}$$

where θ_i indicates the father's income quartile. Figure 6 shows that children from high-income families have higher returns 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 5a).

Figure 6: Returns to experience in different firm classes by father's income



Notes: Estimated returns to firm-class experience from equation (5), by quartile of father's income.

5 The role of skill sorting in the firm pay gradient

The SES gradient in firm pay may partly reflect assortative matching: children from high-income families have higher AKM worker effects, and high-wage workers sort into high-premium firms (e.g., Card et al., 2013). Yet roughly two thirds of the gradient survives conditioning on the estimated worker effects (Table 1, column 5).[29] This "direct" gradient may be genuine, with parental background affecting firm sorting through channels other than skills – information, social and coworker networks, credit constraints – or through preferences for non-pay amenities that vary by family background (Section 6).

But the direct gradient may also be overstated, because the estimated worker effects $\hat{\alpha}_i$ are imperfect proxies of skill. The $\hat{\alpha}_i$ are noisy estimates of α_i, which attenuates the estimated contribution of assortative matching (Dobbin and Zohar, 2025); they capture all *persistent* within-firm earnings differences, not only skills – for example, persistent discrimination against ethnic groups would also load onto α_i;[30] and skills are multidimensional, so conditioning on a uniform skill may miss sorting along the specific dimensions most closely tied to parental background.

We proceed in three steps. First, adding cognitive and non-cognitive skill measures show

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

[30]Family background may also affect earnings *within* firms, in either direction: parents may help their children to be promoted within a firm ("social capital", Dobbin and Zohar, 2025), while "occupational followers" who choose their father's occupation may earn less due to skill mismatch (Bello and Morchio, 2022). Either effect would load onto α_i. Empirically, Stinson and Wignall (2018), San (2022) and Staiger (2026) find that most gains from parental networks arise from access to high-wage firms rather than within-firm advantages.

that skill sorting explains about half of the firm pay gradient at mid-age. Second, over the lifecycle, observed skills explain little of the gradient at labor-market entry but become increasingly important with age. Third, a simple employer-learning model rationalizes these dynamics: employers initially rely on family background as a proxy for skill, and sort on actual skill as information accumulates. Social networks may contribute to the initial pay gaps, but are less consistent with the rising role of skill sorting over the lifecycle.

5.1 Measuring the contribution of skill sorting

To analyze skill sorting more directly, we incorporate explicit measures of cognitive and social (or "non-cognitive") skills from military enlistment tests, and use them alongside education and fixed effects-based skill proxies. These skill measures are highly informative about labor productivity, as reflected in their strong associations with wages and other long-term labor-market outcomes (Lindqvist and Vestman, 2011; Nybom, 2017). Because enlistment testing was mandatory only for men, we restrict this analysis to males.[31]

Specifically, we apply a decomposition in the spirit of Gelbach (2016) to examine sorting on (i) cognitive skills, (ii) social skills, (iii) education and (iv) the estimated AKM worker effect. This framework allows us to test whether the worker effects – often used as controls in worker-firm sorting analyses – are incomplete measures of skill, and whether skill sorting is therefore underestimated. It also lets us compare the relative importance of different skill dimensions – cognitive versus social skills – for firm sorting.

We summarize the decomposition in equations (6a)–(8), suppressing year/age subscripts for simplicity. After estimating the firm pay gradient β_{firm} in equation (6a), we augment this regression with our four mediators of interest, cognitive skill, social skill, education, and the AKM worker effect,

$$\hat{\psi}_j = \mu_\psi + \beta_{firm} y_{f(i)} + \omega_i \tag{6a}$$

$$\hat{\psi}_j = \mu_{\psi,res} + \beta_{firm,direct} y_{f(i)} + \beta_{cog} cog_i + \beta_{soc} social_i + \beta_{edu} edu_i + \beta_{iFE} \hat{\alpha}_i + \upsilon_i \tag{6b}$$

The coefficient $\beta_{firm,direct}$ in equation (6b) thus captures the *direct* effect of family background on firm premia not mediated by observed skills, while the difference $\beta_{firm} - \beta_{firm,direct}$ captures the part explained by the mediators. To quantify how parental income relates to each skill

[31] The military tests are taken at around age 18 and were, in the cohorts we study, compulsory for all men. The cognitive skill score aggregates four subtests measuring verbal, logical, spatial and technical skills. The social/non-cognitive score is based on a half-hour semi-structured interview in which a certified psychologist grades the enlistee both overall and along dimensions such as leadership, social maturity, and emotional stability (Lindqvist and Vestman, 2011). Both scores are standardized to mean zero and unit variance within each birth year. Our education variable is years of education corresponding to one's highest level attained.

measure, we estimate the auxiliary regressions:

$$cog_i = \mu_{cog} + \phi_{cog} y_{f(i)} + \epsilon_{1i} \tag{7a}$$

$$social_i = \mu_{soc} + \phi_{soc} y_{f(i)} + \epsilon_{2i} \tag{7b}$$

$$edu_i = \mu_{edu} + \phi_{edu} y_{f(i)} + \epsilon_{3i} \tag{7c}$$

$$\hat{\alpha}_i = \mu_{iFE} + \phi_{iFE} y_{f(i)} + \epsilon_{4i} \tag{7d}$$

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

$$\beta_{firm,skill} \equiv \beta_{firm} - \beta_{firm,direct} = \beta_{cog}\phi_{cog} + \beta_{soc}\phi_{soc} + \beta_{edu}\phi_{edu} + \beta_{iFE}\phi_{iFE}. \tag{8}$$

Table 4 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), $\hat{\beta}_{firm} = 0.069$, closely matches our baseline estimates for all males (0.070, Table A7) or the full population (0.065, Table 1). 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, consistent with the finding that parental background has a substantial direct influence on firm sorting beyond what is driven by differences in children's skills.[32]

Columns (3)–(6) report the contribution of each of the skill measures to overall sorting, as captured by the $\beta\phi$ terms from equation (8). The estimated worker effects and cognitive skills are key in explaining why children from richer families work at better-paying firms, 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 A11).[33]

We thus conclude that skill sorting is underestimated – and the direct role of family background overstated – when skills are proxied solely by the estimated AKM worker effects. When conditioning only on the worker effect, 72.3% of the SES gradient remains unexplained (0.047/0.065, Table 1), or 71.4% when considering only men (Appendix Table A7).

[32] Figure A18 shows that the firm pay gradient increases markedly along the parental-income distribution (as also visible from Figure 1b), 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 β_{firm}.

[33] One possible interpretation is that the worker effects from the AKM model capture qualitative differences in formal education (e.g., field of study) better than differences in cognitive skills, but that the latter are an important determinant of worker-firm sorting. Note also that education is also strongly correlated with cognitive skills, whereas other skill pairs are more moderately correlated (see Appendix Table A12). We provide further details on the underlying components of the decomposition in Appendix Section A4.1.

Table 4: The contribution of skills to the SES gradient in firm pay at age 40

	Firm pay premia					
	Overall	Unexp.	Cognitive	Social	Education	Worker FE
	β_{firm}	$\beta_{firm,direct}$	$\beta_{cog}\phi_{cog}$	$\beta_{soc}\phi_{soc}$	$\beta_{edu}\phi_{edu}$	$\beta_{iFE}\phi_{iFE}$
	(1)	(2)	(3)	(4)	(5)	(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 β_{firm}	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: Estimates from the decomposition in equations (6a)–(8) using the main sample at age 40, but excluding women and men with missing enlistment scores. Standard errors obtained using 250 bootstrap replications.

Incorporating our full set of skill measures reduces the unexplained share substantially, to 53.6% (Table 4). Still, about half of the firm pay gradient remains unexplained.

5.2 The contribution of skill sorting over the lifecycle

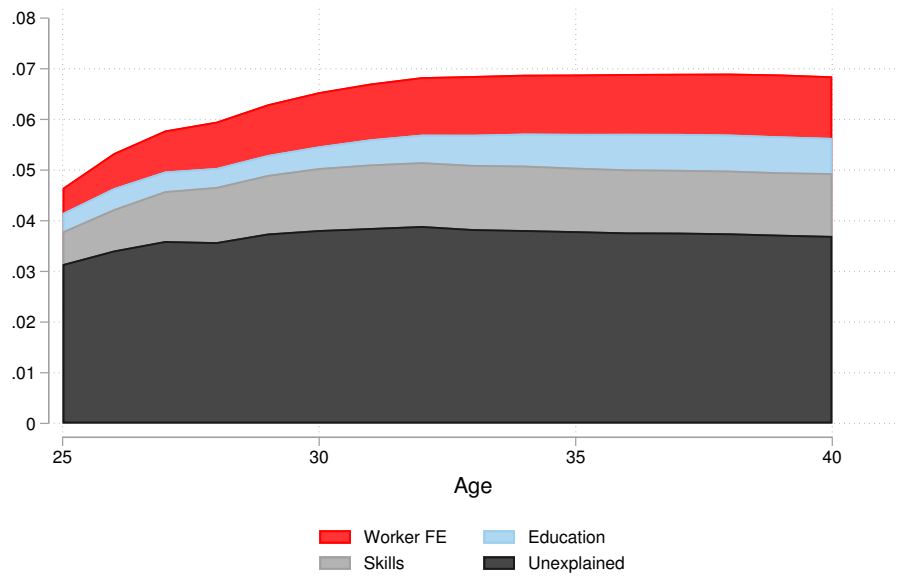
What accounts for the part of the firm pay gradient not explained by skill? One explanation is that family background matters directly, perhaps through parental and peer networks. Alternatively, it may shape sorting more indirectly, through discrimination, or through skills that are correlated with parental income but not well captured by the skill measures in Table 4.[34] Distinguishing between these mechanisms is difficult, but our setting provides some leverage, as beyond rich skill measures, it lets us track skill sorting over the lifecycle – and, as we formalize in Section 5.3, different mechanisms generate different lifecycle patterns.

To trace the role of skill sorting over the lifecycle, we implement the decomposition in equations (6a)–(8) separately by age. Figure 7 shows the results, revealing two key patterns. First, skill sorting accounts for only a small share – about 30% – of the large firm pay gradient β_{firm} at labor market entry; nearly 70% reflects direct family effects not mediated by observed skills. Second, the contribution of skill sorting, $\beta_{firm,skill}$, increases substantially over the lifecycle, while the direct effect of family background not explained by skill, $\beta_{firm,direct}$, remains nearly constant. As a consequence, the relative importance of the two components shifts substantially with age: the share of the gradient attributable to direct family effects falls from nearly 70% at age 25 to about half at age 40.

That only a small share of the firm pay gradient at career start is explained by skill is con-

[34] 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 7: The contribution of skills to the SES gradient in firm pay by age



Notes: Age-specific estimates from the decomposition in equations (6a)–(8) using the main sample, but excluding women and men with missing enlistment scores. The components sum up to the SES gradient in firm premia.

sistent with recent work on the importance of networks: young adults remain more closely tied to their parents and parental networks, and may rely more heavily on family-provided information, referrals, or support. Their own peer networks may depend on parental background as well. However, the second finding – that skills become increasingly predictive of firm sorting as workers age – is harder to reconcile with a pure networks story.

5.3 Firm pay gradients, family background, and imperfect information

To rationalize these lifecycle patterns, we introduce a simple model of skill sorting under imperfect information, in the spirit of statistical discrimination (Phelps, 1972) and employer learning (Farber and Gibbons, 1996; Altonji and Pierret, 2001). Family background affects firm sorting through networks and through its use as a proxy for unobserved skill; the latter channel alone reproduces the patterns in Figure 7.

Setup. Assume that the expected productivity of worker i depends on her skill, $skill_i^*$, which increases in socioeconomic background SES_i^*,

$$skill_i^* = \gamma_{ses} SES_i^* + u_i, \tag{9}$$

where u_i captures idiosyncratic skill components orthogonal to family background, with variances σ^2_{ses} and σ^2_u, and $\gamma_{ses} > 0$ (as confirmed empirically in Appendix Table A11). However,

firms might observe only imperfect proxies of skill and socioeconomic background,

$$skill_i = skill_i^* + v_i, \tag{10}$$

$$SES_i = SES_i^* + w_i, \tag{11}$$

where v_i and w_i represent classical measurement error with variances $\sigma_v^2(a)$ and σ_w^2, respectively, and where $\sigma_v^2(a)$ is indexed by age a as the measurement error in skills may fall as employers observe a worker's performance over time.

Firms infer true skill from these observed proxies for skill and socioeconomic background. Assume further that the conditional expectation $E[skill_i^*|skill_i, SES_i]$ is linear,

$$E[skill_i^*|skill_i, SES_i] = b_{skill}\, skill_i + b_{ses}\, SES_i, \tag{12}$$

where the coefficients b_{skill} and b_{ses} can be derived using the omitted-variable formula (Appendix A6.1). Expressed in terms of the reliability of the two signals,

$$b_{skill} = \frac{\rho_{skill}\,(1 - \pi\,\rho_{ses})}{1 - \pi\,\rho_{skill}\,\rho_{ses}}, \tag{13}$$

$$b_{ses} = \gamma_{ses}\,\frac{\rho_{ses}\,(1 - \rho_{skill})}{1 - \pi\,\rho_{skill}\,\rho_{ses}}, \tag{14}$$

where $\sigma^2_{skill^*} = \gamma^2_{ses}\sigma^2_{ses} + \sigma^2_u$ is the variance of true skill, $\rho_{skill} = \frac{\sigma^2_{skill^*}}{\sigma^2_{skill^*}+\sigma^2_v}$, and $\rho_{ses} = \frac{\sigma^2_{ses}}{\sigma^2_{ses}+\sigma^2_w}$ are the reliability ratios of the skill and background signals (the share of each proxy's variance due to its true component), and $\pi = \frac{\gamma^2_{ses}\sigma^2_{ses}}{\sigma^2_{skill^*}}$ is the share of the skill variance attributable to family background (corresponding to the population R^2 in eq. 9).

Each coefficient rises with the reliability of its own signal and falls with that of the other: the less precisely a firm observes skill, the more weight it places on family background as a substitute proxy. With perfectly observed skill ($\rho_{skill} = 1$), $b_{skill} = 1$ and $b_{ses} = 0$, background carries no information beyond skill.

Finally, we assume that the sorting of workers to firms depends on their productivity and social networks.[35] Specifically, with probability p a worker-firm match is formed through networks that vary with parental background, $\psi_j = \delta_1 SES_i^*$, while with probability $1 - p$ it is formed through expected productivity, $\psi_j = \delta_2\, E[skill_i^*|skill_i, SES_i]$.[36] This implies the

[35]The assumption that workers with higher expected productivity sort into firms with higher pay premia is consistent with standard models of worker-firm sorting (Eeckhout, 2018; Bagger and Lentz, 2019) and the empirical patterns documented by Card et al. (2013) and also found in our data (Appendix Table A13). We approximate this relationship linearly. The assumption that parental and peer networks help high-SES children to secure better jobs is consistent with the evidence from different causal designs, including Kramarz and Skans (2014), Zimmerman (2019), San (2022), Staiger (2026), Eliason et al. (2023) and Campa (2025).

[36]A strand of research points to the use of networks as a device for employers to overcome imperfect information on worker productivity, thus complicating a clear isolation of "network" and "information" effects (Montgomery, 1992; Hensvik and Skans, 2016).

reduced form

$$E[\psi_j|skill_i, SES_i] = \underbrace{\lambda\, SES_i^*}_{\text{network channel}} + \underbrace{\delta\,(b_{skill}\, skill_i + b_{ses}\, SES_i)}_{\text{information channel}}, \tag{15}$$

where $\lambda \equiv p\,\delta_1$ summarizes the influence of parental and peer networks (which depends on both the overall influence of networks and their SES bias) and $\delta \equiv (1-p)\,\delta_2$ summarizes the influence of workers' expected productivity (which depends on both their skill and family background).

Firm pay gradients. This structure allows us to derive model counterparts to the total and direct firm pay gradients defined in equations (6a) and (6b). From eq. (15), regressing firm pay premia ψ_j on true socioeconomic background SES_i^* yields

$$\beta_{firm} = \lambda + \delta\,(b_{skill}\,\gamma_{ses} + b_{ses})\,, \tag{16}$$

while regressing firm pay premia on SES_i^* controlling for $skill_i^*$ implies

$$\beta_{firm,direct} = \lambda + \delta b_{ses} \tag{17}$$

$$\beta_{firm,skill} = \delta\gamma_{ses} b_{skill}. \tag{18}$$

with b_{skill} and b_{ses} defined in equations (13) and (14).

Lifecycle implications. This structure reproduces the central patterns of Figure 7. First, since b_{ses} is proportional to $(1-\rho_{skill})$, the direct gradient $\beta_{firm,direct} = \lambda + \delta\, b_{ses}$ is positive whenever skill is imperfectly observed, and largest when employers know least about it – at career start.

Second, the skill-mediated gradient $\beta_{firm,skill} = \delta\,\gamma_{ses}\, b_{skill}$ increases in ρ_{skill}, and hence over the lifecycle: as firms learn workers' true skill, a growing part of the high-SES advantage is attributed to skill rather than to background as such.

Third, the *overall* firm pay gradient β_{firm} rises over the lifecycle as well. Substituting (13)–(14) into (16),

$$\beta_{firm} = \lambda + \delta\,\gamma_{ses}\,\frac{\rho_{skill} + \rho_{ses} - (1+\pi)\,\rho_{skill}\rho_{ses}}{1 - \pi\,\rho_{skill}\,\rho_{ses}}, \tag{19}$$

which is increasing in ρ_{skill} with $\partial\beta_{firm}/\partial\rho_{skill} \propto 1-\rho_{ses}$ (Appendix A6.1): firms initially underestimate the productivity of high-SES workers if background is imperfectly observed, and revise upward as skills are revealed, widening the SES gradient in firm pay.[37]

The model does not capture one feature of Figure 7: the direct effect $\beta_{firm,direct}$ is nearly

[37] When background is perfectly observed ($\rho_{ses} = 1$), $\beta_{firm} = \delta\gamma_{ses}$ irrespective of ρ_{skill} and there would be no lifecycle dynamic in the overall firm pay gradient; when skill is initially unobserved ($\rho_{skill} \to 0$), $\beta_{firm} = \delta\gamma_{ses}\rho_{ses} = \delta\gamma_{ses}\sigma^2_{ses}/(\sigma^2_{ses} + \sigma^2_w)$.

constant over the lifecycle, whereas the model predicts it to decline as firms learn. Two mechanisms could offset the decline. First, our skill measures are imperfect, so the estimated "direct" effect still contains some skill sorting, which grows as the return to skill rises with age (Appendix A6.1). Second, with search frictions workers may initially work for firms with premia lower than what their skills warrant but climb the job ladder over time, so assortative matching – and hence the effective pass-through δ – rises with age. Because $\beta_{firm,direct} = \delta\, b_{ses}$, such an increase in δ may offset a decline in b_{ses}.

Discussion. To summarize, a model in which firms observe family background and gradually learn workers' skill rationalizes the key lifecycle patterns: a large SES gradient in firm pay at career start that is largely unexplained by skill, its further rise over the lifecycle, and the growing share attributable to skill sorting. None of these requires networks: the model generates such profiles even with the network channel shut down ($\lambda = 0$). Our point is not that networks are unimportant – a growing literature shows that they matter – but that background effects conditional on skill may reflect statistical discrimination as well as networks, which may explain their large magnitude, and that the rising contribution of skill sorting arises naturally under employer learning but is hard to explain by networks alone. The relative importance of the two mechanisms remains an open question.

6 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. Leveraging job-to-job flows in US data, Sorkin (2018) 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 (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 differences in overall compensation (pay plus amenities) across *firms*, however, they need not understate compensation gaps between *groups*, if groups trade off pay and amenities differently. Morchio and Moser (2026), 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. The pay-amenity trade-off may also vary with family background. If children from high-income families feel pressure to match parental earnings and thus select into higher-paying but less pleasant jobs, firm pay gaps would overstate welfare gaps; if affluent backgrounds

instead provide the freedom to choose more enjoyable work, they would understate them.[38]

Whether the SES gradient in firm pay reflects a corresponding gradient in the overall value of jobs is thus ambiguous and has, to our knowledge, not been examined. To make progress, we construct revealed-preference measures of overall firm desirability – or "firm value", denoted V_j – from workers' job-to-job flows, following Sorkin (2018). The firm value captures the full bundle that a firm offers: its pay premium and priced or non-priced amenities, but also attributes such as expected earnings growth and job security.

We find that firm pay premia are informative about overall desirability: high-paying firms poach more workers from other firms, firm values correlate strongly with firm pay, and the SES gradient in firm pay extends to a corresponding gradient in firm value. Sorting into higher-value firms thus makes children from high-income families better off in welfare terms, not only in terms of pay. Moreover, pay gaps understate how their welfare advantage rises with age: while SES gaps in firm pay plateau in the early 30s, gaps in firm value continue to widen throughout our observation window. As with firm pay, conditioning on the estimated worker effects explains about one-third of the value gradient at mid-career.

Amenities may vary across firms for two distinct reasons (Sorkin, 2018): firms offering equal utility may differ in their pay–amenity mix (compensating differentials, or the "Rosen" motive), and firms may differ in the overall utility they offer, forming a job ladder in desirability (the "Mortensen" motive). To assess the role of each margin, we examine SES gradients in firm pay conditional on firm value, and vice versa. While these comparisons have limitations, they show that both margins matter: the firms of high-SES children rank higher in both dimensions of the joint distribution of pay and value.[39] At a given level of overall desirability, their jobs are tilted toward pay rather than amenities – whether reflecting preferences or opportunities – while at a given level of pay, their firms offer more overall value, indicating that their pay premia are less often offset by disamenities.

6.1 SES gradients in firm desirability from job-to-job flows

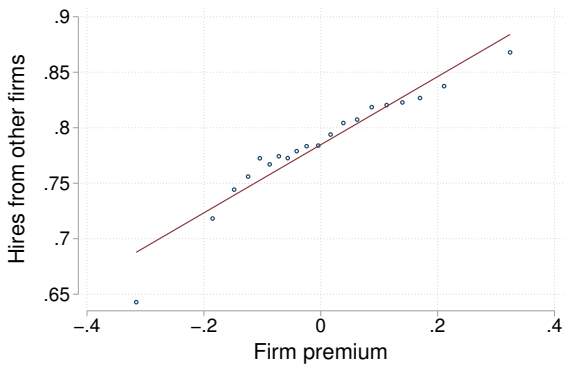
As a starting point, we quantify the extent to which firms "poach" workers from other firms rather than hiring from non-employment – a simple, widely used proxy for firm desirability (Bagger and Lentz, 2019; Crane et al., 2023). Figure 8a shows that high-paying firms (with high $\hat{\psi}_j$) hire predominantly from other firms, while low-paying firms hire more often from non-employment, where workers hold weaker outside options. This pattern aligns with the job ladder implied by standard search models (Burdett and Mortensen, 1998) and confirms

[38] Consistent with the latter hypothesis, Schüle (2026) finds that children from high-SES families place less value on income and job security, and Haeck and Laliberté (2025) show that they are overrepresented in occupations characterized by desirable non-pay attributes.

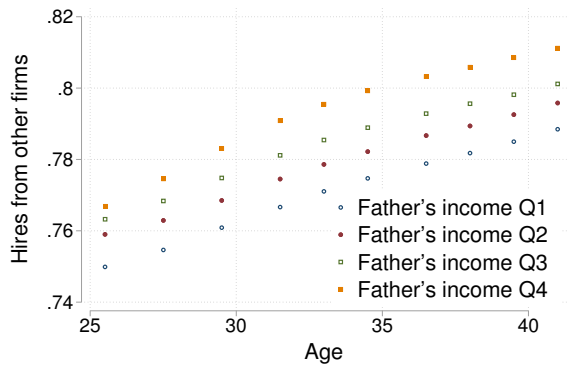
[39] The estimated firm values lack a cardinal scale, so gradients in pay and value cannot be compared in absolute terms, and the overall SES gradient in amenities cannot be signed.

Figure 8: 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 effects based on equation (1). Subfigure (b) shows the mean poaching rate by age and quartile of father's income.

that high-paying firms tend to be more attractive to workers.

Figure 8b shows that the poaching rate of the current employer rises with parental income and increases more steeply over age for children from the top quartile, mirroring the pattern in firm pay premia. Yet while average firm pay peaks in the early 30s before declining slightly (cf. Figure 2), poaching rates keep rising throughout our age window. These SES gaps in poaching close when conditioning on firm pay (Appendix Figure A19), a first indication that pay gaps are reasonable proxies for SES gradients in broader firm value.[40]

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 voluntary job-to-job transitions to infer workers' ordinal preferences over firms. The idea is that if workers voluntarily go from one firm to another, it must mean they prefer the new firm over the previous one.[41] The implied firm value V_j captures the overall value of employment at firm j, consisting of both the firm's wage premium and amenities.[42]

A first important observation is that 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 A15),

[40] 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 4a).

[41] We define voluntary employer-to-employer transitions as moves that do not involve unemployment benefit receipt or zero-earnings spells between consecutive firm spells; see further details on our implementation in Appendix A5.1.

[42] Firm pay and firm values vary substantially across industries (see Figure A20). For example, the education industry exhibits high firm values relative to its wage component, whereas oil and gas extraction offers the reverse, consistent with compensating differentials. Other industries align on both dimensions: finance ranks high and the restaurant industry ranks low in both pay and value. Men and workers with more education and higher skills are more likely to work in high-value firms (Table A14).

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, and consistent with our findings for poaching rates, firm pay premia are informative about overall firm desirability.

To draw more direct inference about differences in welfare, we examine SES gradients in firm value. Table 5 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 mean zero and to have the same standard deviation as $\hat{\psi}_j$.

We first confirm that the SES gradient in firm pay (column 1) extends to a corresponding gradient in overall firm value (column 3), and thus the overall desirability of firms.[43] Sorting into higher-value employers therefore makes children from high-SES families better off. While the gradients cannot be compared in absolute terms – the firm value lacks an interpretable scale – the R^2 is scale-invariant and permits a relative comparison. Parental income explains a smaller share of the variation in firm value than in firm pay (R^2 of 0.019 vs. 0.031), a gap that is not driven by sampling error.[44] Still, the SES gradient in firm value is more than twice as large as one would expect if high-SES children sorted purely on pay – sorting does not operate on pay alone.[45]

Figure 9 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. As careers progress, all workers, but particularly high-SES workers, increasingly sort into firms offering better amenities or other non-pay benefits. Although the levels of the two gradients cannot be compared, this contrast in age profiles is informative: focusing on firm pay alone would miss that the sorting of high-SES children into more desirable firms keeps intensifying beyond the early 30s.

As shown in Panel B of Table 5, controlling for the estimated worker fixed effects reduces the mid-career SES gradients in firm pay and firm value by a similar proportion – about one-third (cf. Panels A and B, columns 1 and 3). Skill sorting thus partly explains why high-SES workers sort into more desirable firms, and it does so to a similar extent for firm values as for firm pay premia, supporting a common interpretation of the two gradients. Yet, even

[43] 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). In addition, because we use information on unemployment benefit receipt to distinguish voluntary employer-to-employer transitions, the estimation period is restricted to 1993–2018. The sample size is thus slightly reduced. However, 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).

[44] Firm values contain greater sampling error, but a split-sample correction reduces the gap only marginally; we return to the role of sampling error below.

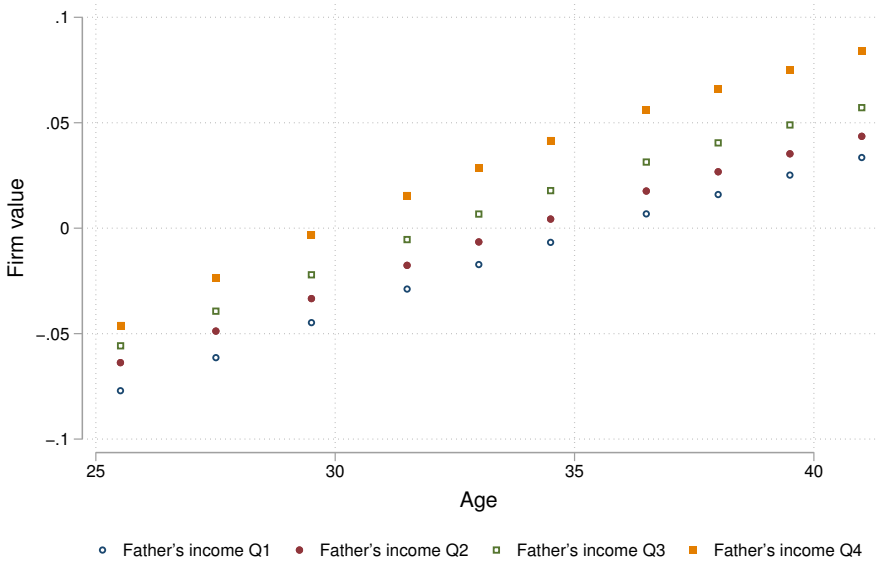
[45] Given the correlation of 0.52 between firm pay and firm value, sorting purely on pay would imply an R^2 in firm value of only $0.52^2 \times 0.031 \approx 0.008$, falling well short of the observed R^2 in firm value.

Table 5: 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	0.041	0.014
	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{V}_j$		0.588		
		(0.001)		
$\hat{\psi}_j$				0.446
				(0.001)
R^2	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	0.028	0.009
	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{V}_j$		0.549		
		(0.001)		
$\hat{\psi}_j$				0.422
				(0.001)
$\hat{\alpha}_i$	0.152	0.088	0.117	0.053
	(0.001)	(0.001)	(0.001)	(0.001)
R^2	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	0.040	0.013
	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{V}_j$		0.635		
		(0.001)		
$\hat{\psi}_j$				0.474
				(0.001)
R^2	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 premium $\hat{\psi}_j$ 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 footnote 50 in Section A6.2 for details. The estimated firm values $\hat{V}_j$ are standardized to have mean zero and the same standard deviation as the firm pay estimates. Robust standard errors in parentheses.

Figure 9: Firm values over the lifecycle



Notes: Mean firm value over the lifecycle, by quartile of father's income. The firm value is estimated following Sorkin (2018).

conditional on skills, high-SES children remain considerably more likely to work in higher-value firms.

6.2 Pay or amenities? Interpreting conditional SES gradients

The firm value captures everything that makes an employer attractive: the pay premium and amenities, but also forward-looking components such as expected earnings growth and job security. Without further assumptions, we cannot distinguish the relative roles of these components. Nor do our results indicate whether high-SES children receive better or worse amenities in absolute terms – only that the value they derive from higher firm pay is larger than any compensated disamenities. With these caveats in mind, we examine SES gradients in firm pay conditional on firm value, and vice versa – asking whether the pay advantage of high-SES children merely reflects the overall desirability of their employers, and whether their employers are more desirable than their pay alone would suggest.

In column (2) of Table 5, we regress firm pay on both parental income and the estimated firm value. The coefficient on parental income falls by about 40% relative to column (1), but remains positive: conditional on working at similarly valued firms, children from high-SES backgrounds work at firms that offer higher pay – and thus fewer (priced) non-pay benefits. 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. Panel A, 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.[46]

A first concern is that sampling error in the estimated firm values and pay premia would mechanically generate such positive conditional gradients: noise attenuates the coefficient on the conditioning variable, and the SES gradient it would otherwise absorb spills over to parental income. In Appendix A6.2, we formalize this argument and address it with a split-sample IV strategy, instrumenting each estimate with its counterpart from an independent subsample. As shown in Panel C, the conditional SES gradients decline only modestly (by about 10%) and remain positive and substantial; they are also robust to excluding small firms, for which the estimated firm values are noisier (Appendix Table A17). Sampling error therefore does not drive these patterns, though specification error arising from the model's assumptions cannot be ruled out.

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). Read jointly, however, the two estimates simply indicate that the firms of high-SES children rank higher in both dimensions of the joint distribution of pay and value. Such a pattern requires two motives for amenity variation: differences in the pay–amenity mix among firms of similar value (compensating differentials, or the "Rosen" motive), and *sufficiently large* dispersion in overall desirability across firms (the "Mortensen" motive) – with high-SES children advantaged on both. It would arise, for instance, 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.[47]

What do these patterns imply about amenities? The R^2 comparison above could not distinguish whether amenities are only weakly related to parental background or partly offset pay differences (compensating differentials); the conditional gradients point to elements of both. On the one hand, column (2) implies that, at a given level of overall desirability, high-SES children hold jobs tilted toward pay rather than amenities – consistent with a stronger weight on pay, whether this reflects their preferences or the opportunities open to them.[48] On the other hand, column (4) implies that their pay premia are not primarily compensation for disamenities: at a given level of pay, their firms offer more overall value. At the

[46] When we additionally control for the estimated AKM worker effect (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.

[47] 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.

[48] Consistent with this interpretation, the correlation between firm pay and firm value is slightly higher among high-SES children (Figure A21 and Table A16). Within-group correlations should be interpreted with caution, however, since selective sorting restricts the range of pay and values within groups.

same pay premium, the firms of low-SES children thus more often carry offsetting disamenities, whereas the pay premia of high-SES children more often constitute "rents" that raise the overall value of their jobs.[49] Part of the conditional value advantage may, moreover, reflect dynamic rather than non-wage benefits: firm values are forward-looking, and Section 4 showed that high-SES children sort into firms characterized by higher wage growth. What we cannot determine is whether the overall SES gradient in amenities is positive or negative: this would require expressing firm values in pay units, which their ordinal scale does not permit.

Taken together, the evidence suggests that firm pay premia, while an imperfect proxy for overall firm desirability, capture meaningful socioeconomic differences in job quality. High-SES children work in firms that rank higher in both pay and non-pay dimensions, implying that the SES gradient in firm pay reflects a genuine gradient in the overall quality of jobs.

7 Conclusions

The contribution of worker-firm sorting to intergenerational earnings persistence appears remarkably strong: in our data, static firm pay premia account for 36% of the intergenerational elasticity of earnings around age 40, rising to about one half once dynamic firm-specific returns to experience are taken into account. These pay differences do not merely reflect compensating differentials. Indeed, conditional on pay, children from high-income families work at more desirable firms, and this advantage continues to grow in midlife even as their pay advantage has already stabilized.

At first glance, such a large role for firms may seem to collide with the literature's traditional emphasis on human capital. For the most part, however, it does not: much of the firm pay gradient reflects skill sorting, as children from high-income backgrounds sort into higher-paying firms precisely because they enter the labor market with better skills and more education. Worker-firm sorting thus amplifies skill gradients that emerge before labor market entry – it is a key mechanism by which skill differences translate into earnings differences.

However, the large SES gradient in firm pay that arises already at career start is *not* well explained by skills, but appears consistent with recent evidence on the role of parental and peer networks in labor market entry. The subsequent widening of the firm pay gradient, by contrast, is almost entirely accounted for by skills, and is thus better explained by employer learning under imperfect information. Such informational frictions would also contribute to the unexplained pay gaps at career start, perhaps explaining why these gaps are so large.

These results link the nascent literature on worker-firm sorting and intergenerational per-

[49] This finding is in line with evidence by Haeck and Laliberté (2025), who show that low-SES children in Canada are overrepresented in physically demanding industries such as mining, oil and gas, or construction. We find a similar pattern in Sweden (Appendix Figure A20). Further, Stansbury and Rodriguez (2026) argue that differences in preferences cannot explain class gaps in career advancement within academia.

sistence to the literature on peer effects in labor market access, but more work is required. Networks and skill sorting under imperfect information need not be distinct mechanisms, as connections may serve precisely to overcome informational frictions about worker skill. How such informational frictions shape worker-firm sorting has, however, received little attention in the intergenerational literature, and strikes us as a promising avenue for future research.

From a policy perspective, measures that improve the skill acquisition of children from lower-income families thus have a double payoff: they not only raise earnings directly – the "worker component" of pay – but also improve sorting into better-paying firms. Our findings suggest that this second payoff may be as important as the first for intergenerational mobility. However, the large pay gaps at career start deserve particular attention. Social mixing might not suffice to break the influence of networks (Einiö, 2026). And if firms use family background as a proxy for unobserved skills, then what matters is not only the skills that children from different backgrounds acquire, but also how visible these skills are to employers.

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 (2026): "Differences in On-the-Job Learning Across Firms," *Journal of Labor Economics*, 44, 149–188.

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 of 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. Jäntti, 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 (2026): "How do neighborhoods and firms affect intergenerational mobility?" Census CES Working Paper 26-18, U.S. Census Bureau, Center for Economic Studies.

CATTAN, S., K. G. SALVANES, AND E. TOMINEY (2025): "First-generation elite: The role of school social networks," *American Economic Review*, 115, 4369–4403.

CHETTY, R., M. O. JACKSON, T. KUCHLER, J. STROEBEL, N. HENDREN, R. B. FLUEGGE, S. GONG, F. GONZALEZ, A. GRONDIN, M. JACOB, D. JOHNSTON,

M. KOENEN, E. LAGUNA-MUGGENBURG, F. MUDEKEREZA, T. RUTTER, N. THOR, W. TOWNSEND, R. ZHANG, M. BAILEY, P. BARBERÁ, M. BHOLE, AND N. WERNERFELT (2022): "Social capital I: measurement and associations with economic mobility," *Nature*, 608, 108–121.

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 (2025): "Quantifying the Role of Firms in Intergenerational Mobility," *American Economic Journal: Applied Economics*, forthcoming.

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.

EINIÖ, E. (2026): "The Long-Term Impacts of Mixing the Rich and Poor: Evidence from Conscript Dorms," *American Economic Journal: Applied Economics*, 18, 177–209.

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*, 88, 600–626.

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.

FISCHER, A., A. GORSHKOV, T. M. SANDOY, AND J. WALLDORF (2025): "Peers, Heirs, and Careers: Labor Market Effects of Alumni Networks," Tech. rep., Working paper.

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*, 19, 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.

HENSVIK, L. AND O. N. SKANS (2016): "Social networks, employee selection, and labor market outcomes," *Journal of Labor Economics*, 34, 825–867.

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.

LALIBERTÉ, J.-W. AND A. WHALLEY (2026): "Social Connections and the Persistence of Income Across Generations," Tech. rep., IZA Discussion Paper No. 18691.

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 (2021): "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. (2003): *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.

MONTGOMERY, J. D. (1992): "Job search and network composition: Implications of the strength-of-weak-ties hypothesis," *American Sociological Review*, 57, 586–596.

MORCHIO, I. AND C. MOSER (2026): "The Gender Pay Gap: Micro Sources and Macro Consequences," *American Economic Review*, 116, 1765–1810.

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, 239–268.

PHELPS, E. S. (1972): "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62, 659–661.

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

PUTNAM, R. D. (1995): "Bowling alone: America's declining social capital," *Journal of democracy*, 6, 65–78.

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?" CESifo Working Paper 9842, CESifo.

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. (2026): "The Intergenerational Transmission of Employers and the Earnings of Young Workers," *Review of Economics and Statistics*, forthcoming.

STANSBURY, A. AND K. RODRIGUEZ (2026): "The Class Gap in Career Progression: Evidence From U.S. Academia," *Econometrica*, 94, 1345–1373.

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.

WILLIS, S. (2024): "It's Who You Used to Know: Professional Networks, Heterogeneity, and Inequality," Tech. rep., Working Paper, Uppsala University.

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

Appendix

Labor-market Drivers of Intergenerational Earnings Persistence

A1 Additional results

A1.1 Sample construction and descriptive statistics

Table A1 reports summary statistics for the AKM sample and the main intergenerational sample. Differences between the samples 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, 11 years earlier, resulting in lower college attainment and slightly lower log earnings. The average firm size is roughly 450 employees.

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 A1: Descriptive statistics

	AKM sample	Main sample (born 1967–77)	
Age	20–64	25–41	39–41
Earnings years	1985–2018	1992–2018	2006–2018
Earnings	301,558	316,664	400,878
Share women	0.49	0.48	0.48
Share with college degree	0.39	0.48	0.50
Year of birth	1961	1972	1972
Firm size	455.89	435.90	460.56
Number of 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. Earnings are measured as annual earnings in Swedish kronor (SEK), expressed in 2018 prices.

Table A2: 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:			
Worker 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.

A1.2 Variance decomposition of the AKM model

Using our estimates from equation (1), we can decompose the variance in log earnings 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 A2, 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 worker 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 30–38% across the three samples. Firm fixed effects explain between 8–13% of the variance in log earnings, while the covariance between firm and worker fixed effects explains another 7–12%. 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 0.40 ($0.019/(\sqrt{0.101}\sqrt{0.022})$).

Differences in firm pay and worker-firm sorting together explain 15–24% of the total

Table A3: 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%	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 log earnings. Columns (2)–(4) report the slope coefficients from the corresponding regressions when decomposing child log earnings y_{ijt} according to equation (1) into worker fixed effects α_i, mean firm fixed effects ψ_j for ages 39–41, and time-varying controls. Robust standard errors in parentheses.

variance, depending on sample. For the main sample observed over the lifecycle (column 2) we find a somewhat decreased importance of the worker component, compared to the full AKM sample (column 1). For the prime-age version of the main sample (column 3), which only includes earnings 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).

A1.3 Decomposition of the IGE

A1.3.1 Alternative measures of parental income

In our baseline, family background is measured as father's log earnings. Table A3 reports the decomposition of the IGE when parental income is instead measured as the average of the father's and mother's log earnings. The overall IGE is then larger (0.323), which almost exclusively stems from a much lower variance in the average compared to the father-based parental income measure (Pearson correlations are roughly similar). However, the firm pay gradient – the coefficient on parental income when the dependent variable is the AKM firm effect (column 3) – remains proportionally similar either way, at 0.103 or 32% of the IGE.

A1.3.2 Measurement error and the AKM estimation window

Our baseline AKM model is estimated over a long panel (1985–2018), whereas related studies rely on much shorter windows (six years in Dobbin and Zohar, 2025; five-year windows

in Card et al., 2026). To examine how the length of the estimation window affects our results, we re-estimate the model over shorter periods: once over the years 2006–2018, the period in which the children in our sample (born 1967–1977) are observed at ages 39–41, and separately over the five-year window centered on the year in which each birth cohort turns 40 (2005–2009 for the 1967 cohort, up to 2014–2018 for the 1976 cohort). For each birth cohort, we then estimate the SES gradient in firm pay by regressing the mean estimated firm premium at ages 39–41 on log father's income. Figure A1 plots the resulting cohort-specific gradients, based on firm premia from the 13-year window and from the cohort-specific five-year windows, and Table A4 reports worker-weighted averages across cohorts. The underlying samples differ, as the connected set is smaller in short windows and more firms fall below the minimum of five movers; subfigure (a) uses all available firm premia, while subfigure (b) restricts both series to the same set of workers and firms.

The figure documents two findings. First, the SES gradient in firm pay is fairly constant across birth cohorts, in particular when the firm premia are estimated over the longer window. Second, the estimated gradient is systematically smaller when the premia are estimated over the short windows – by 0.006 log points on average across cohorts, or about 15%, and by 0.005 (12%) when holding the sample fixed (Table A4, columns 1 and 2) – and it fluctuates more across cohorts. Because sample composition is held fixed in subfigure (b), this remaining gap is not explained by the loss of small firms in shorter windows.

Classical measurement error in the estimated premia cannot account for this pattern, as the premia serve as the dependent variable in these regressions. Figure A2 suggests that the deviations between short- and long-window estimates are instead mean-reverting: plotting the deviation against the long-window premium (in binned means, weighted at the worker level), short-window premia exceed the long-window estimates for firms with low premia and fall below them for firms with high premia. Part of this negative relation is mechanical, as estimation error in the long-window premia on the horizontal axis itself induces mean reversion, but together with the smaller SES gradients it suggests that short-window premia are compressed in a way that is correlated with parental income.

Two mechanisms can explain why shorter windows and less representative samples yield smaller firm pay gradients. The first is composition: some of the firms with the most extreme pay premia are small, and children from low-income families are more likely to work in small firms with particularly low premia. Such firms are the first to drop out of the connected set when the data are poorer or the estimation window is shorter, so that the gradient falls even if each premium were measured without error – as it does when we raise the minimum number of movers per firm (Table A6, Panels D–F). The second mechanism is specific to short windows: annual earnings in a year in which a worker changes employer combine pay at the origin and the destination firm but are attributed to only one of them, which pulls the estimated premia of high- and low-paying firms toward each other. Even when earnings are

reported separately by employer, the transition-year observation covers only part of the year, unless earnings are expressed per month or partial spells are excluded. Whether this matters therefore depends on the data structure – with monthly or quarterly records and reliable spell dates, earnings can be attributed to the correct employer and partial periods dropped (see Engbom et al., 2023) – but it is a potential issue in most settings based on annual earnings. In a long window, most observations lie in the interior of a worker-firm spell, so that transition years matter little; in a five-year window they make up a much larger share of the observations that identify each firm's premium.

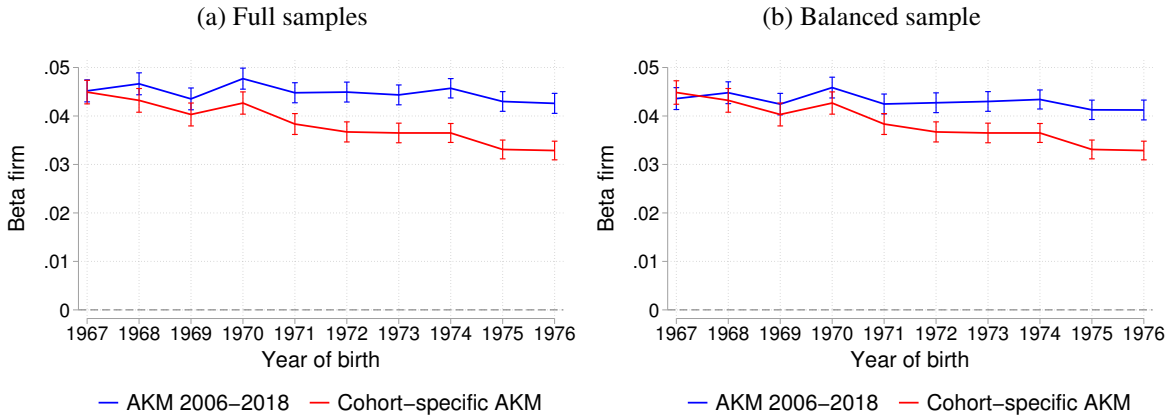
Figure A3 and the right-hand columns of Table A4 examine the second mechanism directly, re-estimating both sets of premia after dropping the final year at the origin firm and the first year at the destination firm. Dropping these years removes short spells and thus many small firms from the connected set, which lowers the gradient for both windows – the composition effect described above, illustrated by the lower estimate in column (3) compared to those in columns (1) and (2). But once transition years are removed and the sample is held fixed, short- and long-window estimates are close (0.030 versus 0.032, a deviation of less than 10%). The shortfall of short-window estimates thus largely reflects the contamination of transition years, while the remaining sensitivity to the estimation window operates through the coverage of small firms.

Two further observations support this reading. The compression is not an artifact of imposing time-invariant premia over a long panel: allowing premia to vary across four subperiods of 1985–2018, estimated jointly so that stayers connect adjacent windows (Table A6, Panel A), modestly raises the estimated firm contribution. And randomly dropping half of the workers before the AKM estimation – which reduces precision without shortening the panel or altering the set of firms – leaves the results unchanged (Table A5). The contribution of firms to intergenerational persistence thus appears understated when the AKM model is estimated over short panels or on samples with limited coverage of small firms, one reason why we find a larger firm share than studies relying on shorter panels.

A1.3.3 Alternative specifications for the AKM regression

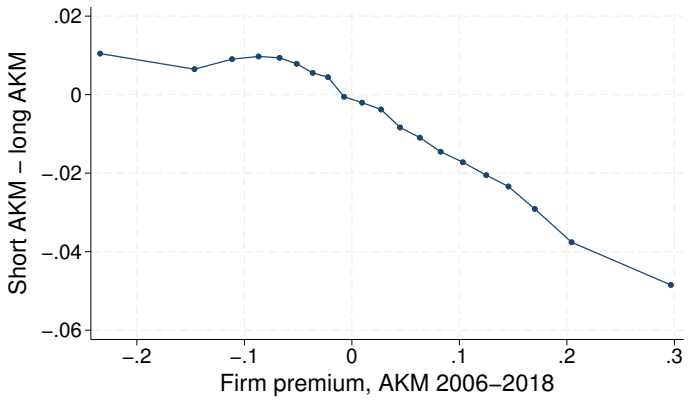
Table A6 reports the decomposition of the IGE under alternative AKM specifications. Panel A allows firm premia to vary over time by splitting the 1985–2018 period into four subperiods, which modestly increases the estimated role of firms. Note that in this exercise we estimate the premia for all windows jointly, analogous to Engbom et al. (2023): worker fixed effects remain constant across windows, and each worker effectively switches firm identifiers at the window boundaries, so that stayers connect the premia of adjacent windows. In Figure A1 we instead estimate a separate AKM regression for each window, so that neither channel links the estimates across windows. Panel B defines pay premia at the establishment rather

Figure A1: The SES gradient in firm pay by birth cohort and AKM estimation window



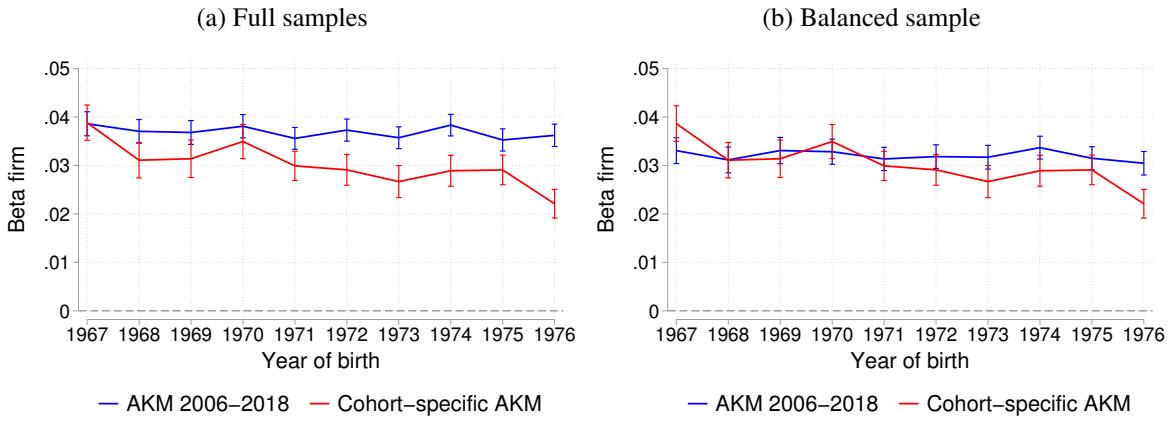
Notes: Cohort-specific estimates of the SES gradient in firm pay, from regressions of the mean estimated firm premium at ages 39–41 on log father's income, for children born 1967–1976. The first series (blue lines) uses firm premia from an AKM model estimated over the years 2006–2018; the second series (red line) uses premia from separate AKM models estimated over the five-year window, centered around the year in which the respective cohort is aged 40 (2005–2009, ..., 2014–2018). Subfigure (a) uses all available firm premia in each specification, while subfigure (b) restricts both series to the same set of workers. Vertical bars indicate 95% confidence intervals.

Figure A2: Firm premia estimated over short vs. long panels



Notes: Binned scatter plot comparing firm premia estimated over short and long panels. The horizontal axis shows the firm premium from an AKM model estimated over the years 2006–2018; the vertical axis shows the deviation of the premium estimated over the cohort-specific five-year windows (see Figure A1) from this long-window estimate. Observations are weighted at the worker level, with one observation per child in our sample, based on the firm of employment at ages 39–41.

Figure A3: The SES gradient in firm pay by birth cohort and AKM estimation window, transition years removed



Notes: Cohort-specific estimates of the SES gradient in firm pay, from regressions of the mean estimated firm premium at ages 39–41 on log father's income, for children born 1967–1976. The AKM models are estimated after removing transition years, defined as the first and last year working in a firm. The first series (blue lines) uses firm premia from an AKM model estimated over the years 2006–2018; the second series (red line) uses premia from separate AKM models estimated over the five-year window, centered around the year in which the respective cohort is aged 40 (2005–2009, ..., 2014–2018). Subfigure (a) uses all available firm premia in each specification, while subfigure (b) restricts both series to the same set of workers. Vertical bars indicate 95% confidence intervals.

Table A4: The SES gradient in firm pay, with long and short AKM window

	Baseline:			Transition years removed:	
Samples	Separate	Balanced	Col. (5) sample	Separate	Balanced
	(1)	(2)	(3)	(4)	(5)
Long AKM	0.045	0.043	0.037	0.037	0.032
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Short AKM	0.038	0.038	0.032	0.030	0.030
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Long - Short	0.006	0.005	0.004	0.007	0.002
Number of firms long	151,686	123,484	33,559	86,408	33,559
Number of firms short	123,774	123,484	33,559	33,624	33,559
Number of workers long	749,361	697,446	367,702	580,062	367,702
Number of workers short	698,034	697,446	367,702	368,221	367,702

Notes: Worker-weighted averages of the estimated SES gradient in firm pay across the 1967–1976 birth cohorts. Worker-weighted standard deviations across cohorts are shown in parentheses. Long AKM refers to firm effects estimated over 2006–2018. Short AKM refers to cohort-specific five-year AKM estimates, centered around the year the cohort is age 40. Long – short is the difference between the worker-weighted mean coefficients. Transition-year removed specifications exclude the final year at the origin firm and the first year at the destination firm. Balanced-sample specifications require both long and short AKM firm effects to be observed.

Table A5: 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	100%	60%	31%	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	100%	59%	31%	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 worker fixed effects α_i and mean firm fixed effects ψ_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.

than the firm level, raising the firm contribution slightly relative to our baseline definition, while always using firm identifiers regardless of firm size (Panel C) lowers it, largely due to public-sector employment. Panels D–F vary the minimum number of movers required per firm: requiring at least 10 or 50 movers reduces the firm pay gradient relative to retaining all firms, reflecting the trade-off between limited-mobility bias and the representativeness of the firm sample.

A1.3.4 Decomposition of the IGE: Heterogeneity

Table A7 shows the decomposition of the intergenerational earnings elasticity for different subsamples. Panels A and B split the sample by gender: the overall IGE is larger for men (0.229) than for women (0.165), but the firm pay gradient in column (3) – 0.070 and 0.059, respectively – accounts for a similar share of the IGE in both groups. Panel C excludes children who work in the same firm as their father; the firm pay gradient is nearly unchanged (0.063), indicating that direct parental workplaces account for little of the overall gradient.

Table A6: Decomposition of the IGE using alternative AKM specifications

	Dependent variable				
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_{j=J(i,t)}$	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}} + \hat{\varepsilon}_{ijt}$	$\hat{\psi}_{j=J(i,t)}\vert\hat{\alpha}_i$
	(1)	(2)	(3)	(4)	(5)
A. AKM with time-varying firm FE as in Engbom et al. (2023)					
$y_{f(i)}$	0.198	0.112	0.071	0.015	0.053
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	100%	57%	36%	8%	27%
Worker obs.	826,314	826,314	826,314	826,314	826,314
B. AKM estimated using establishment identifiers					
$y_{f(i)}$	0.198	0.113	0.068	0.018	0.050
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	100%	57%	34%	9%	25%
Worker obs.	831,883	831,883	831,883	831,883	831,883
C. AKM estimated using firm identifiers					
$y_{f(i)}$	0.198	0.128	0.052	0.019	0.033
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Share of IGE	100%	65%	26%	10%	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	100%	57%	35%	9%	27%
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	100%	59%	31%	9%	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	100%	62%	28%	10%	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 for all firms when estimating equation (1) while in Panel C we use firm codes. Our baseline specification uses establishment codes only for multi-establishment firms with more than 1,000 employees (average over 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 A7: Heterogeneity in the IGE decomposition

	Dependent variable				
	y_{ijt}	$\hat{\alpha}_i$	$\hat{\psi}_{j=J(i,t)}$	$\boldsymbol{X}_{it}\hat{\boldsymbol{\beta}}+\hat{\varepsilon}_{ijt}$	$\hat{\psi}_{j=J(i,t)}\vert\hat{\alpha}_i$
	(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	100%	60%	31%	9%	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	100%	57%	30%	13%	24%
Worker obs.	408,498	408,498	408,498	408,498	408,498
			C. Excluding workers who work in the same firm as their 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	100%	59%	32%	9%	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	100%	54%	36%	10%	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 worker fixed effects α_i, mean of firm fixed effects ψ_j for ages 39–41, and time-varying controls. Panel A shows results for males, and panel B shows results for women. In Panel C, workers who work in the same firm as their fathers are excluded, where working in the same firms as the father is defined as having ever worked in the father's main firm (main firm is the firm the father works in for most years between 1985–2018). Panel D shows results where public sector firms are excluded, where public sector firms are defined as firms in industries with SNI92 codes 75, 80, 85, 90, 91, 92, 93 and 99.

Panel D excludes public-sector firms, which raises the firm pay gradient to 0.082, consistent with more compressed wage structures and homogeneous pay-setting policies in the public sector.

A1.3.5 Adding controls

Table A8 examines the role of occupations in the subsample with occupation information. The unconditional firm pay gradient in this sample is 0.055; adding controls for region, indus-

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	0.042	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
Firm size					X		X
Occ. FEs						X	X
Share of β_{firm}	100%	35%	36%	11%	24%	42%	82%
Worker obs.	288,389	288,389	288,389	288,389	288,389	288,389	288,389

Notes: Column (1) reports estimates of the slope coefficient from regressing $\hat{\psi}_j$, at age 40 as estimated from equation (1), on log father's 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), or firm size (10 deciles), occupation fixed effects (2-digit level), or adding all controls simultaneously. Robust standard errors in parentheses.

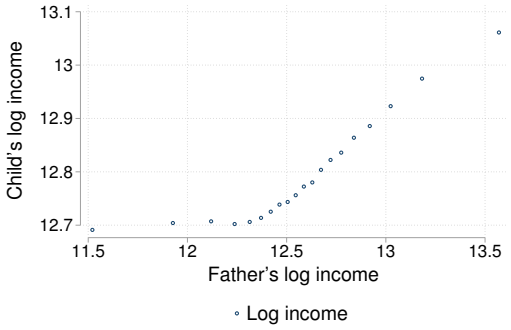
try, public-sector status and firm size, and additionally for occupation fixed effects, reduces it to 0.010 in the fullest specification. Sorting into occupations thus accounts for an important part of the sorting into higher-paying firms, over and above these other characteristics.

A1.4 Non-linear firm pay gradients

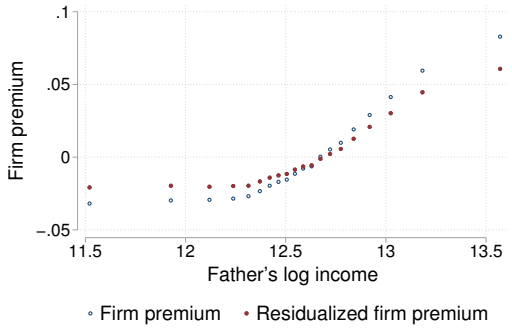
Figure A4 complements Section 2.1 by plotting child log income and firm premia against father's log income.

Figure A4: Child income and firm premia by father's income (logs)

(a) Child log income by father's log income



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



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

A2 Additional evidence on lifecycle dynamics

This appendix provides additional evidence on the lifecycle patterns documented in Section 3. Panels (a)–(c) of Figure A5 plot mean firm pay premia by potential experience rather than age, for the full sample and by education group. Figure A6 shows the firm pay profile conditional on worker fixed effects, Figure A7 separates the profiles by gender, and Figure A8 plots the composition of coworkers – their worker fixed effects and the share with fathers in the top quartile – and the five-year earnings growth of coworkers who stay in the firm and of all coworkers.

A2.1 Firm switching

Figures A9, A10 and A11 disaggregate the firm-switching patterns of Section 3.2 by education group, and Figure A12 separates voluntary from involuntary switches.

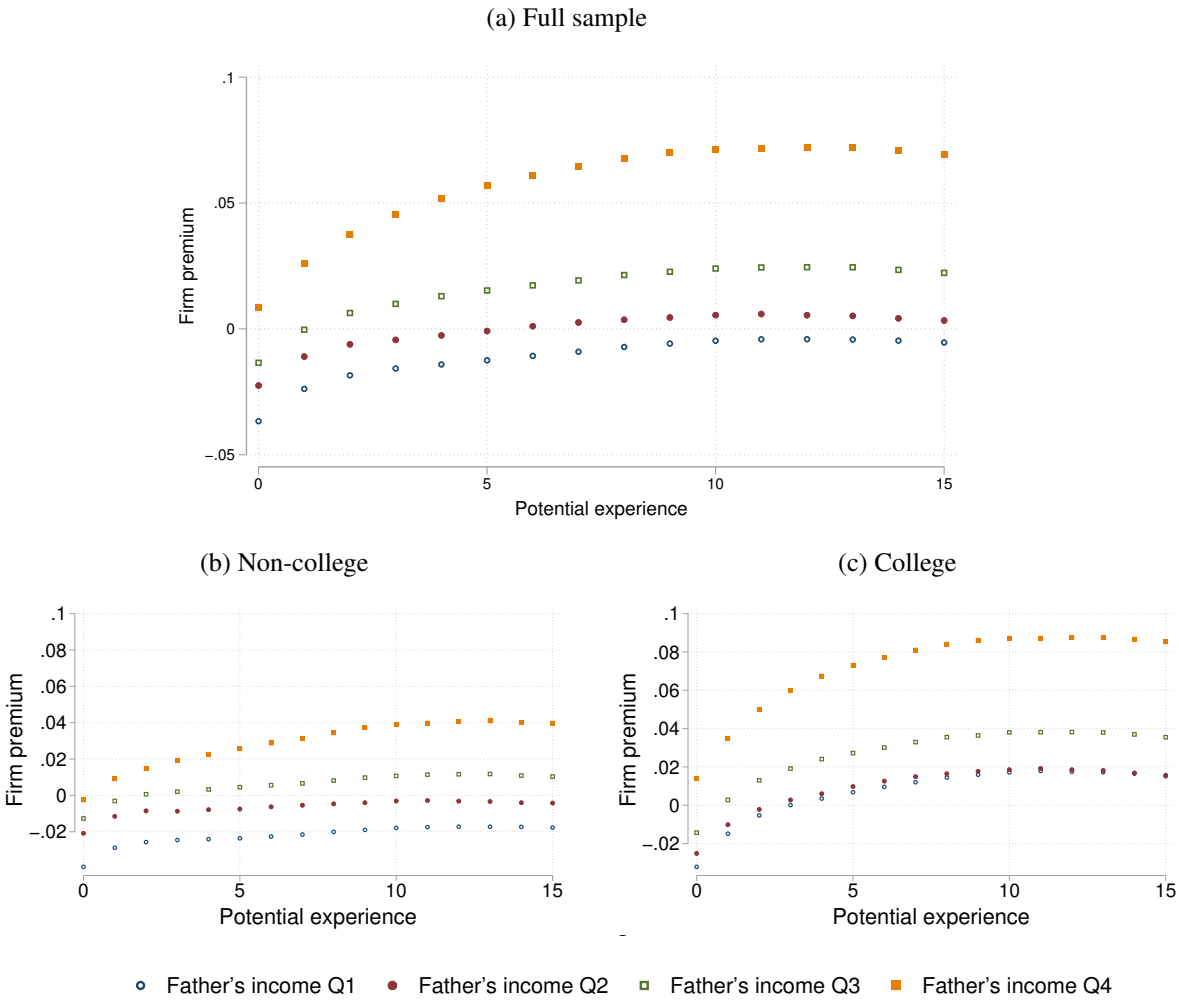
A2.2 Commuting and working patterns

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, the probability of working in the same firm is much higher for children from high-income fathers than for those from low-income families.

One potential explanation for the more frequent job switching and larger gains in firm pay among high-SES children is greater geographic 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., 2021; Agrawal et al., 2024). At the same time, commuting is costly, so firm pay differences associated with commuting may overstate welfare differences.

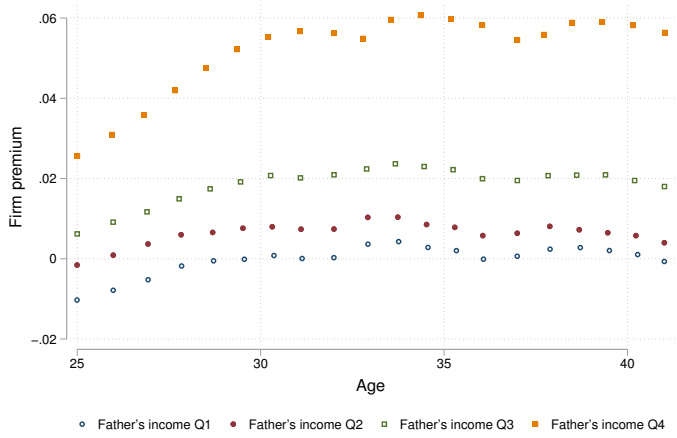
Consistent with this mechanism, Figure A13a shows that children from high-income families are substantially more likely to commute (defined as working in a different municipality than their residence). At age 25, the commuting rate is 8 percentage points higher in the top than in the bottom parental-income quartile, rising to 10 points by age 40. These patterns

Figure A5: Firm pay premia by potential experience



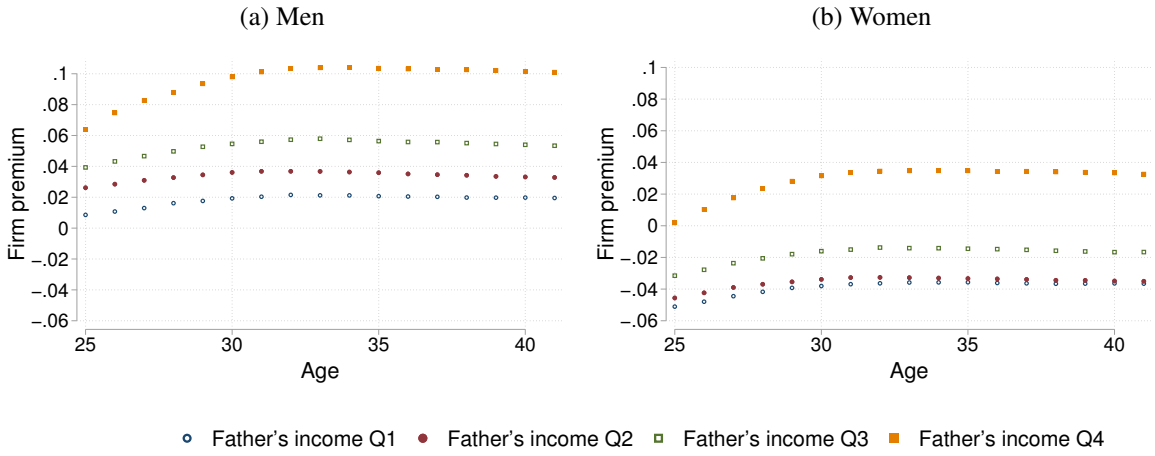
Notes: Mean estimated firm premium $\hat{\psi}_j$ by potential experience, by quartile of father's income. Potential experience is defined as the calendar year minus the year the individual enters the labor market. Labor-market entry is defined as the first year after age 20 in which annual earnings exceed the low-earnings threshold, conditional on having reached the potential age of school completion. The low-earnings threshold is defined as 20% of the median annual earnings of 45-year-old men, and the potential age of school completion as years of education plus six. Subfigure (a) shows the results for the full sample, (b) for children without college education, and (c) for children with college education. Fathers' income quartiles are defined in the full sample.

Figure A6: Firm pay premium over the lifecycle conditional on worker fixed effects



Notes: Mean estimated firm premium $\hat{\psi}_j$, residualized on the worker fixed effects, over the lifecycle and by quartile of father's income.

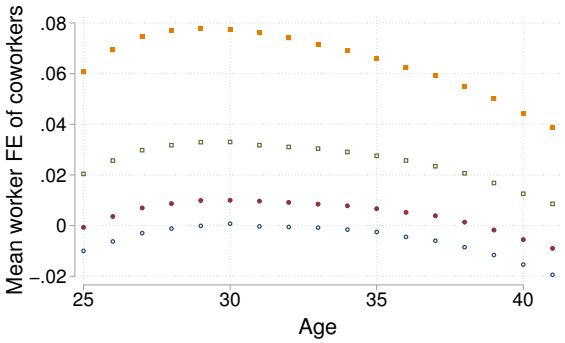
Figure A7: Firm pay premium over the lifecycle by gender



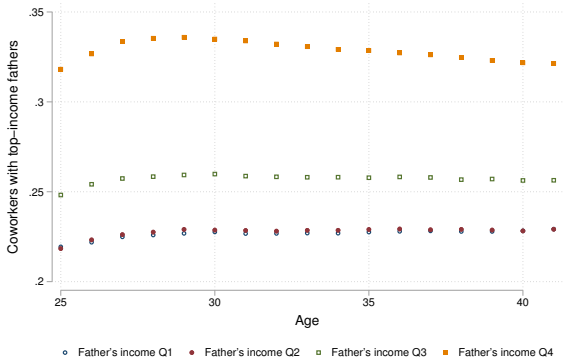
Notes: Mean estimated firm premium $\hat{\psi}_j$ over the lifecycle, by quartile of father's income. Subfigure (a) shows the results for men and (b) shows the results for women. Fathers' income quartiles are defined in the full sample.

Figure A8: Coworker composition and earnings growth

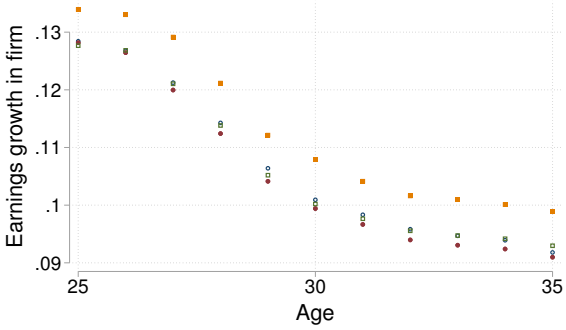
(a) Worker fixed effects of coworkers



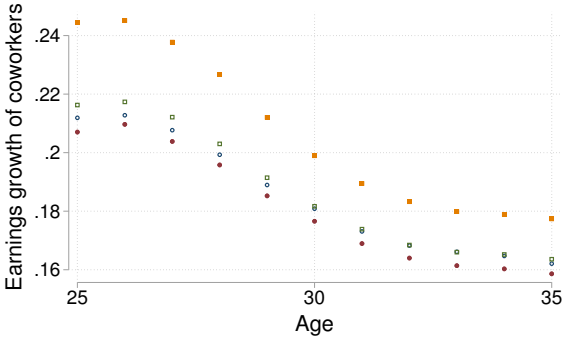
(b) Share of coworkers with fathers in the top parental-income quartile



(c) Five-year earnings growth of coworkers (stayers)



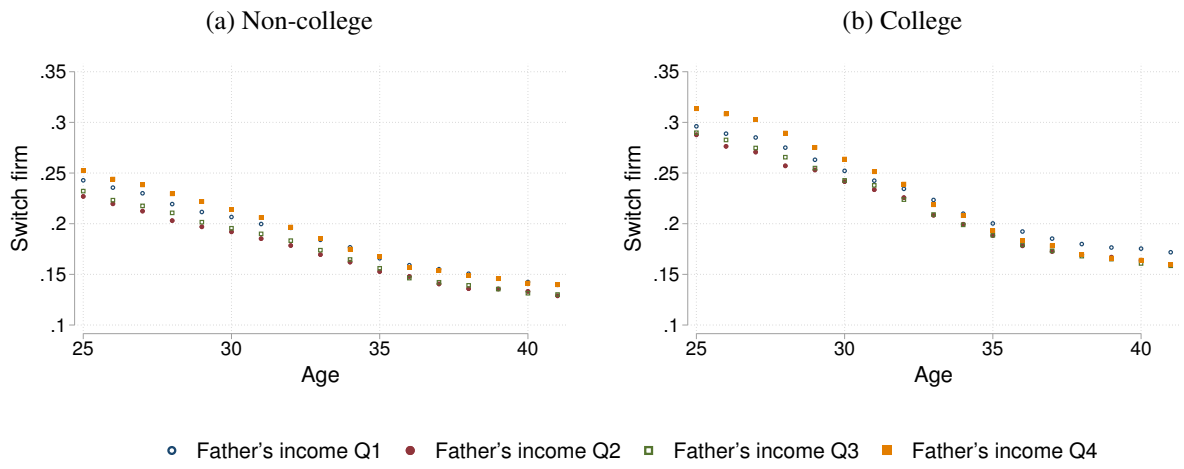
(d) Five-year earnings growth of coworkers (all)



Father's income Q1 Father's income Q2 Father's income Q3 Father's income Q4

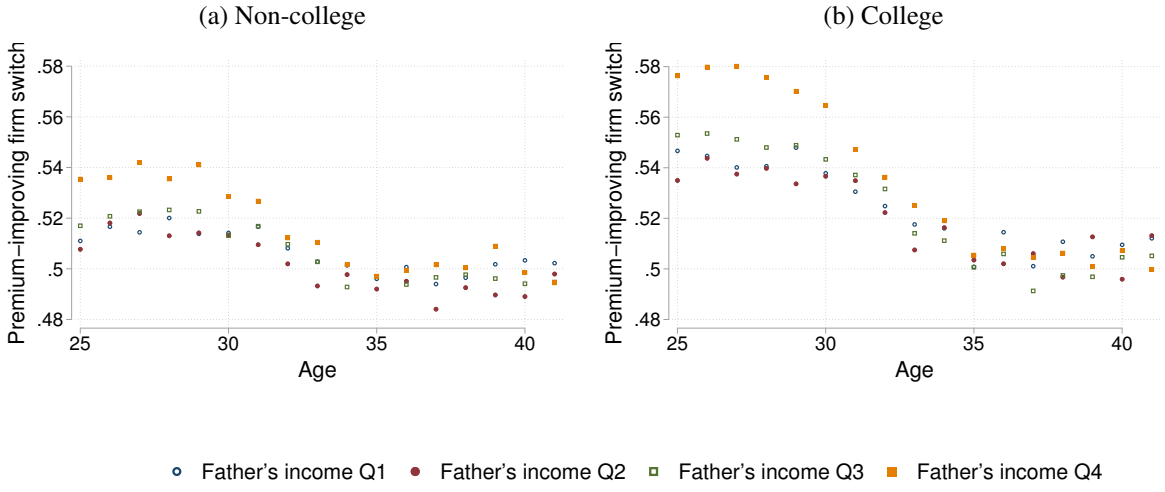
Notes: Subfigure (a) shows the mean worker fixed effect of coworkers and subfigure (b) the share of coworkers with fathers in the top income quartile, by quartile of own father's income. Subfigures (c) and (d) show the mean five-year earnings growth (from year t to $t+5$) of coworkers who stay in the firm and of all coworkers, respectively; subfigure (d) reproduces Figure 4b.

Figure A9: Firm switching over the lifecycle



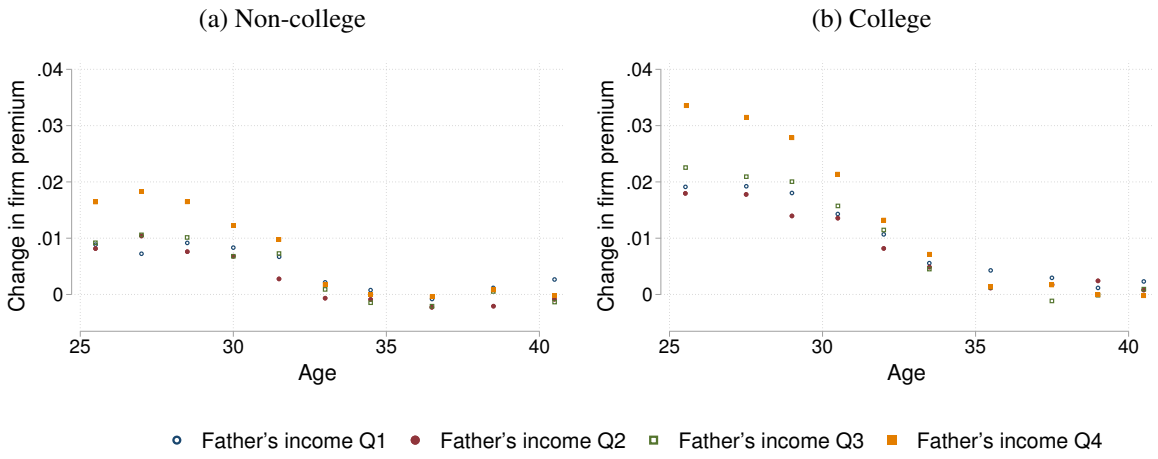
Notes: Patterns of firm switching over the lifecycle, by quartile of father's income. Subfigure (a) shows the results for children without college education, and (b) for children with college education.

Figure A10: Proportion of premium-improving firm switches



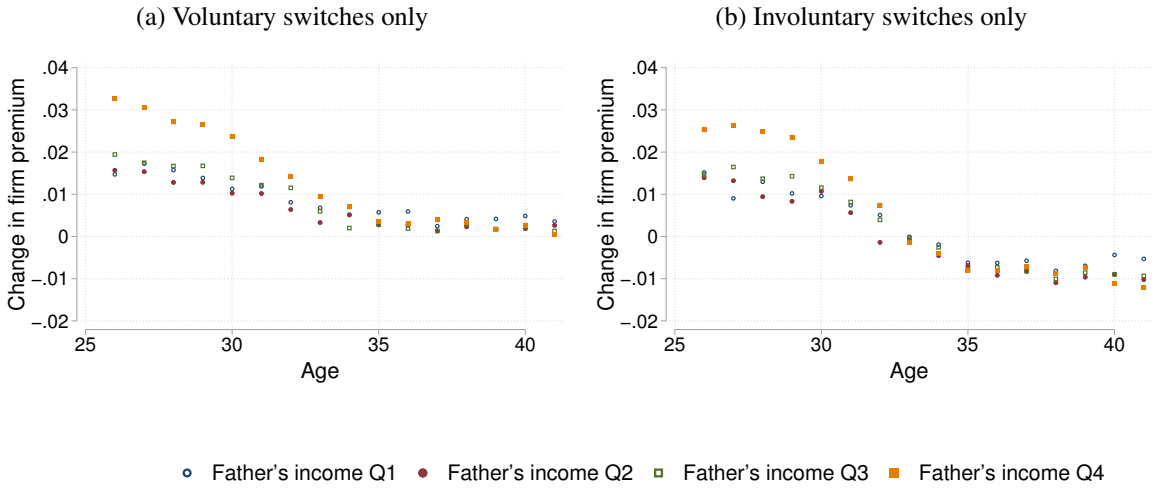
Notes: Probability of switching to a firm with a higher firm premium than the previous one, conditional on switching, by father's income quartile. Subfigure (a) shows the results for children without college education, and (b) for children with college education.

Figure A11: Mean change in firm premium among switchers



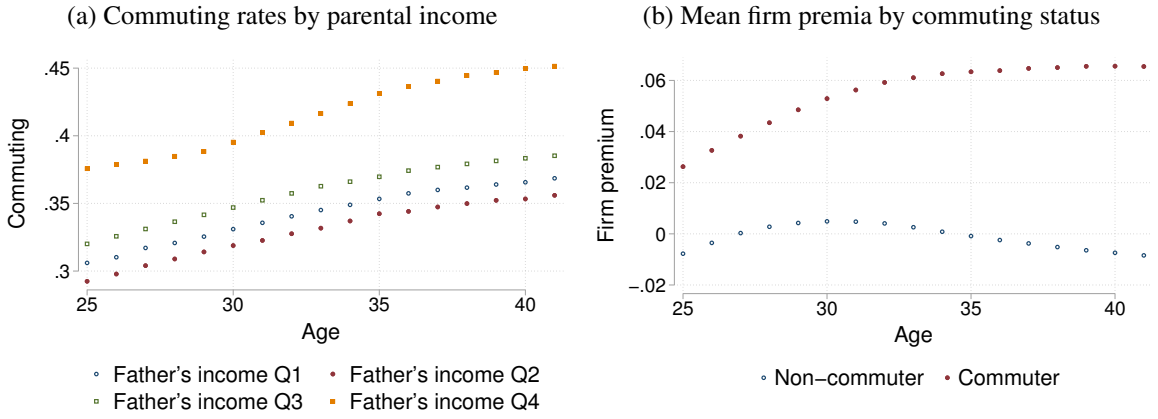
Notes: Difference between the new firm premium and the previous firm premium for individuals who switch firms. Subfigure (a) shows the results for children without college education, and (b) for children with college education.

Figure A12: Change in firm pay premium: voluntary vs. involuntary switches



Notes: Difference between the new firm premium and the previous firm premium for individuals who switch firms, by father's income quartile. Subfigure (a) shows the results for voluntary switches, defined as a switch without any intermediate receipt of unemployment insurance and without any intermediate year of zero income. Subfigure (b) shows the results 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.

Figure A13: Commuting patterns



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

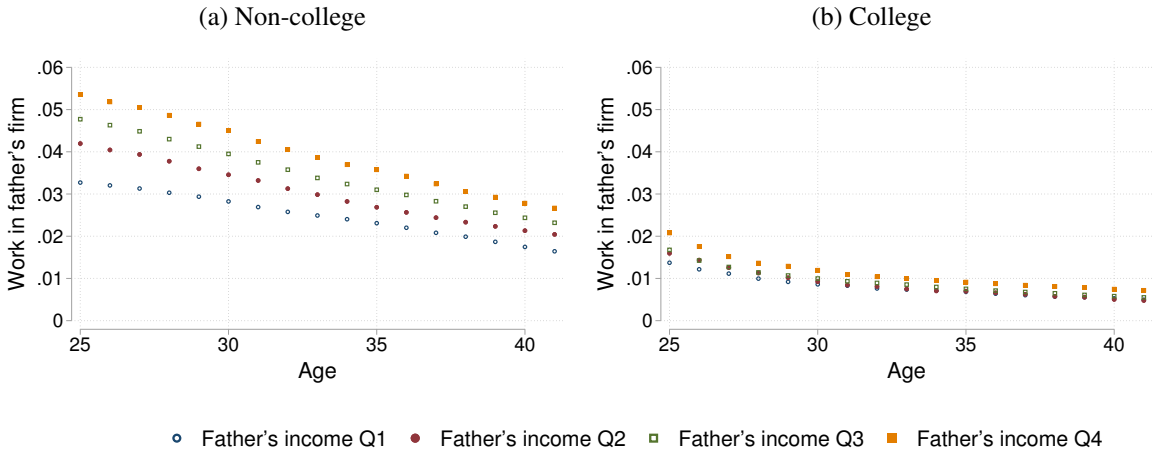
persist within education groups (Appendix Figure A15), indicating that education alone does not explain the SES gradient in commuting.

Figure A13b 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, nearly half a standard deviation. Much of this difference, however, reflects selection. Event-study estimates conditioning on worker fixed effects (Appendix Table A9) 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.

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 each education group, validating the pattern from Figure A13a.

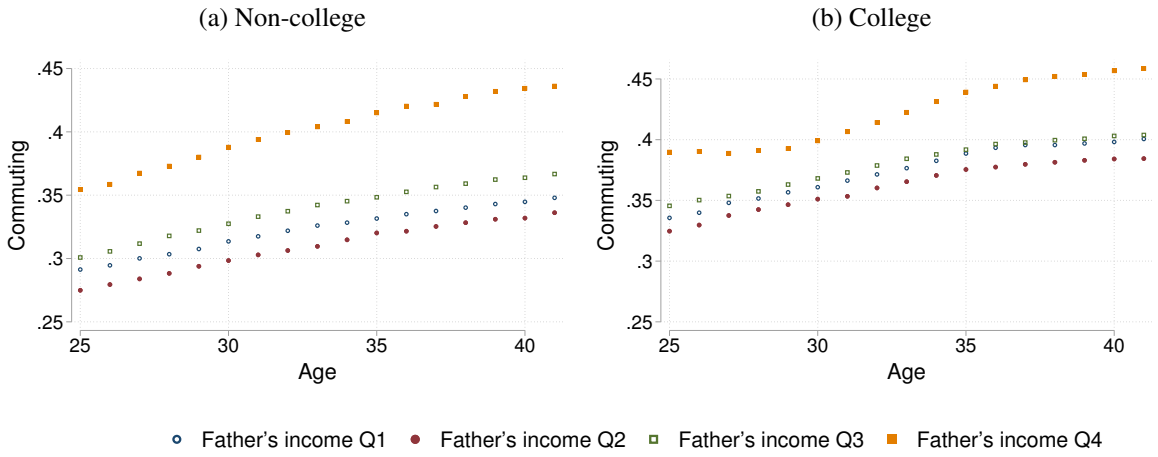
Finally, Table A9 provides event-study-type regression results conditional on worker fixed effects, showing how firm pay changes for individuals who begin to commute. While the firm premium increases when individuals start to commute, this increase is much smaller than the raw wage gap between commuters and non-commuters in the cross-section as shown in Figure A13. As shown in column (1), commuting raises firm premia by about 1 percentage point. 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: Share working in the same firm as their father



Notes: Proportion of children who work in their father's main firm at different ages, separately by quartile of father's income. The father's main firm is defined as the firm in which the father works for the most years between the years 1985–2018. Subfigure (a) shows the results for children without college education, and (b) for children with college education.

Figure A15: Commuting rates over the lifecycle



Notes: Proportion of individuals who commute (i.e., work in a municipality other than the one in which they live) over the lifecycle, by quartile of father's income. Subfigure (a) shows the results for children without college education, and (b) for children with college education.

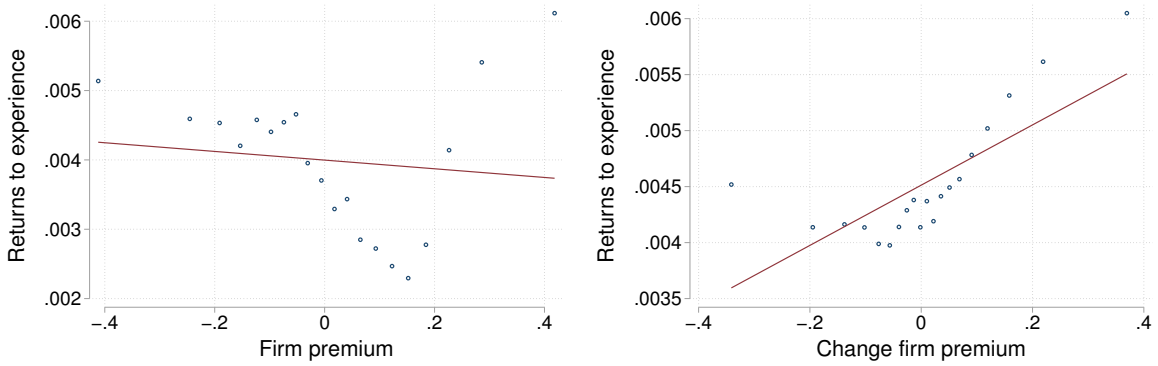
Table A9: Commuting

	Dependent variable: $\hat{\psi}_{j=J(i,t)}$		
	(1)	(2)	(3)
Commuting indicator	0.014	0.022	0.014
	(0.000)	(0.000)	(0.000)
Commuting × 2nd father's income quartile			-0.001
			(0.000)
Commuting × 3rd father's income quartile			0.001
			(0.000)
Commuting × 4th father's income quartile			-0.001
			(0.000)
Commuting × age normalized at 40		0.001	
		(0.000)	
Age controls	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes
Observations	12,977,062	12,977,062	12,977,062

Notes: Regression of estimated firm pay premia on an indicator for commuting (working and residing in different municipalities), worker fixed effects, and controls. Columns (1) and (3) include flexible age controls (age, age squared, age interacted with father's income quartile, and age squared interacted with father's income quartile). Column (2) includes age terms linearly and interacted with commuting. To ease interpretation, age is normalized to age 40 in column 2. Column (3) interacts the commuting indicator with the father's income quartile.

Figure A16: Firm fixed effects and returns to experience

(a) Firm fixed effects and returns to experience (b) Change in firm fixed effects and returns to experience

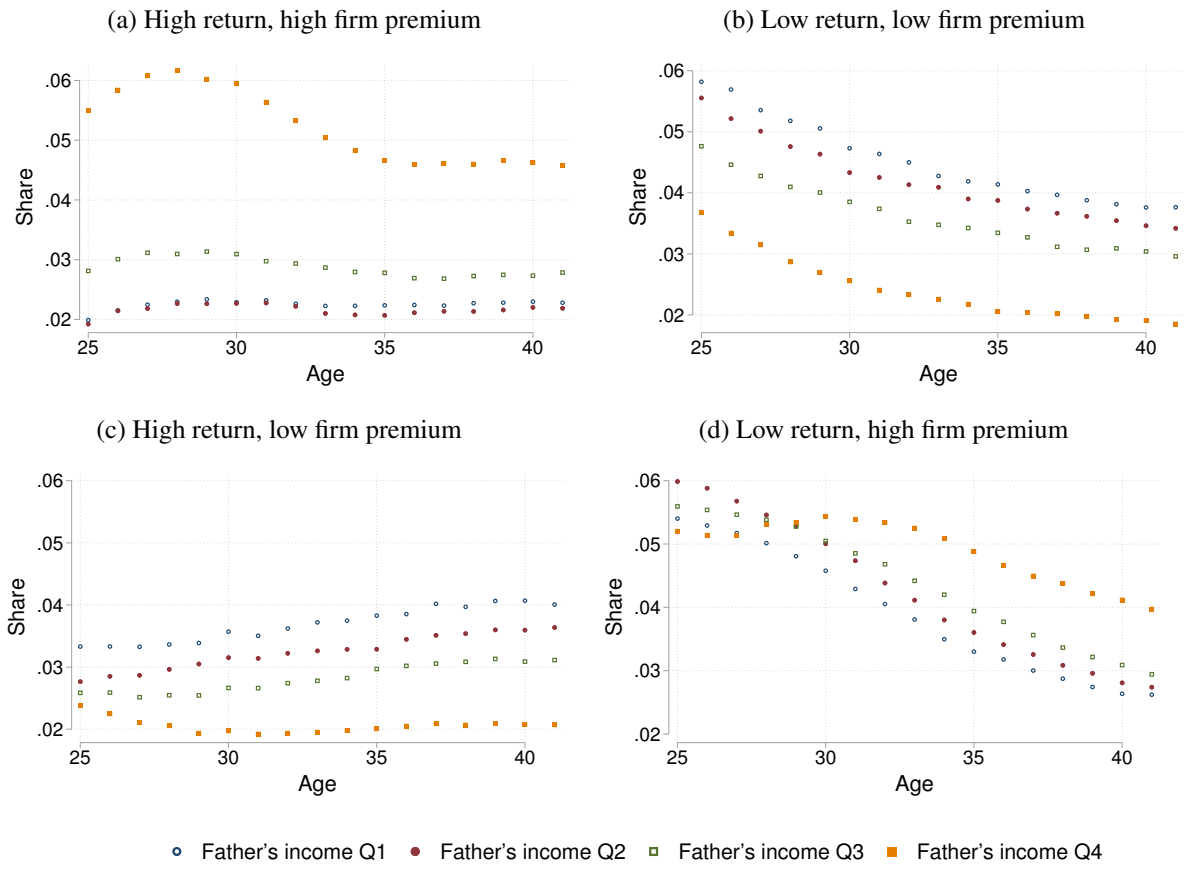


Notes: Firm fixed effects and returns to firm-specific experience estimated with regression (4). Subfigure (a) shows the relationship between returns to experience in the firm class and firm fixed effects. Subfigure (b) shows the relationship between the change in firm fixed effects for individuals who change firms, and the return to experience in the firm class the individuals switch from.

A3 Return to experience in firm classes

This appendix complements the analysis of firm-specific returns to experience in Section 4. Figures A16a and A16b relate the estimated returns to experience of firm classes to their pay premia and to changes in premia upon switching. The relationship between firm-class returns and average pay premia is strongly non-monotonic. This is not surprising given how the classes are constructed: as in Arellano-Bover and Saltiel (2026), firms are classified by the within-firm distribution of earnings growth rather than its mean, so classes with similar average returns can contain quite different firms – and thus very different average pay premia. Arellano-Bover and Saltiel (2026) similarly find a non-monotonic relationship, and overall only a weakly positive correlation, between class-specific returns and average firm pay premia. Figure A17 shows examples of firms with high and low returns and high and low pay premia, and Table A10 decomposes the relationship between returns to experience and parental income.

Figure A17: Share working in firms with high or low pay premia and high or low returns to experience



Notes: Different combinations of working in firms with high/low returns to experience and high/low firm fixed effects. High (low) returns are defined as firms in the two groups with the highest (lowest) returns, and high (low) firm premium is defined as a firm fixed effect in the top (bottom) 20 percent.

Table A10: Decomposing the relationship between returns to experience and parental income

	Dependent variable: Estimated returns to experience					
	(1)	(2)	(3)	(4)	(5)	(6)
$y_{f(i)}$	0.056	0.032	0.042	0.058	0.056	0.028
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Region FEs		X				X
Industry FEs			X			X
Public sector				X		X
Firm size					X	X
Share at age 40	100%	43%	25%	-3%	1%	51%
R^2	0.002	0.016	0.075	0.015	0.008	0.083
Worker obs.	368,913	368,913	368,913	368,913	368,913	368,913

Notes: Column (1) reports estimates of the slope coefficient from regressing the return to experience in the firm class the individual works in at age 40, as estimated from equation (4), on the father's log earnings. Father's log earnings are divided by 100 to improve the readability of the coefficient estimates. 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 firm size (10 deciles). Robust standard errors in parentheses.

Table A11: Skills and father's income

	Dependent variable			
	Cognitive skills	Social skills	Education	$\hat{\alpha}_i$
	(1)	(2)	(3)	(4)
$y_{f(i)}$	1.019	0.713	1.478	0.131
	(0.007)	(0.006)	(0.009)	(0.001)
Observations	361,165	361,165	361,165	361,165
R^2	0.059	0.038	0.070	0.053

Notes: Estimates of $\phi_{cog}, \phi_{soc}, \phi_{edu}$ and ϕ_{iFE} in the Gelbach decomposition; see equations (7a)–(7d). The columns show separate regressions of each of the skill measures on log father's income.

A4 Skill sorting

A4.1 Skills and parental income

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 A11 reports estimates from the auxiliary regressions (7a)–(7d) of each skill measure on parental income (i.e., estimates of $\phi_{cog}, \phi_{soc}, \phi_{edu}$ and ϕ_{iFE}). 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 moderate share to the firm pay gradient (see Table 4). The comparatively low correlation for our proxy of social skills is possibly explained by measurement error, as social skill measures tend to be noisier than measures of cognitive skills (Grönqvist et al., 2017).

Table A13 reports estimates of regression (6b), showing that the coefficients change only marginally when excluding parental income from the regression.

A4.2 Skill sorting across the parental-income distribution

Figure A18 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 (6a)–(8) at each decile of parental income. Figure A18 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

Table A12: Correlations between the different skill measures

	Cognitive skills	Social skills	Education	$\hat{\alpha}_i$
	(1)	(2)	(3)	(4)
Cognitive skills	1.000			
Social skills	0.366	1.000		
Education	0.534	0.301	1.000	
$\hat{\alpha}_i$	0.340	0.362	0.357	1.000

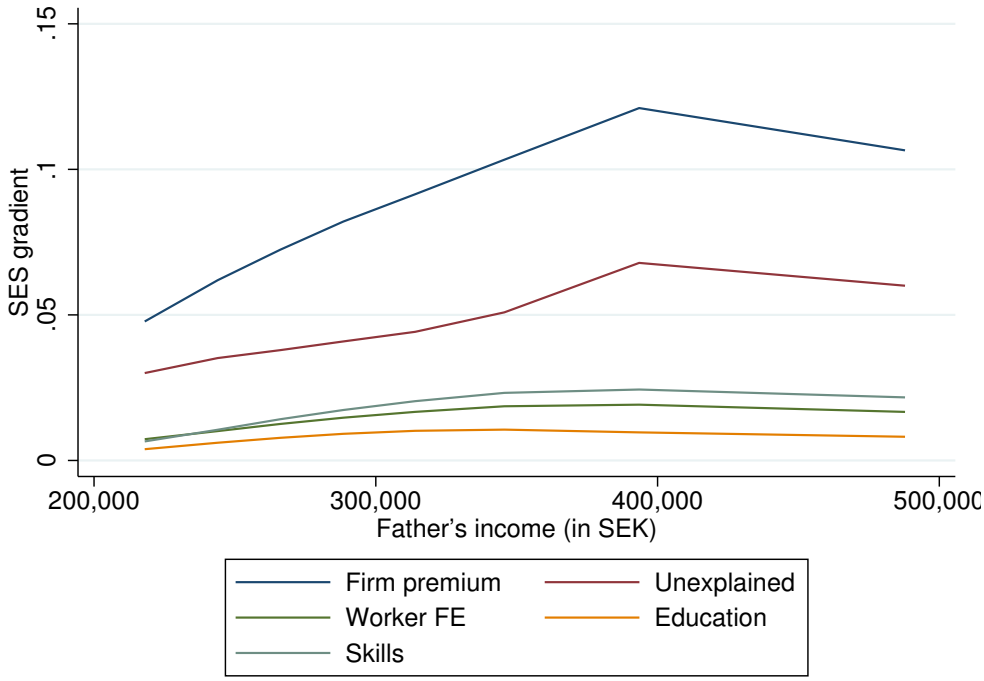
Notes: Pairwise correlations between the different skill measures.

Table A13: Firm pay premia and skills

	Dependent variable	
	$\hat{\psi}_{j=J(i,t)}$	$\hat{\psi}_{j=J(i,t)}$
	(1)	(2)
$y_{f(i)}$	0.037	
	(0.001)	
Cognitive skills	0.011	0.012
	(0.000)	(0.000)
Social skills	0.002	0.002
	(0.000)	(0.000)
Education	0.005	0.006
	(0.000)	(0.000)
$\hat{\alpha}_i$	0.093	0.101
	(0.001)	(0.001)
Observations	361,165	361,165
R^2	0.105	0.095

Notes: Results for the β-coefficients from the Gelbach decomposition. Column (1) shows the results from equation (6b), regressing the firm premium on father's log income and each of the skill measures. Column (2) shows the same regression but excluding father's log income.

Figure A18: Local linear decomposition of skill sorting over the income distribution



Notes: Local linear regressions, from the decomposition in equations (6a)–(8) using the main sample, but excluding women and men with missing enlistment scores. Local linear regressions around the mean income at each decile of parental income, using a bandwidth of 50,000 Swedish kronor and an Epanechnikov kernel. Firm premium reports the SES gradient in firm pay premia, ,$\hat{\beta}_{firm}$. Unexplained reports the remaining SES gradient after accounting for the mediators, $\hat{\beta}_{firm,\ direct}$. Worker FE, education and skills report the respective contributions of these mediators to the SES gradient in firm pay premia.

of β_{firm}. Both skills (driven by the cognitive measure) and the worker 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.

A5 Firm values and compensating differentials

A5.1 Estimating firm values

We here provide information about the estimation of the firm values and some additional results. The firm values are estimated using the method and code from Sorkin (2018). Below is a summary; for a more detailed description of the estimation see Sorkin (2018).

The underlying logic of the firm-value estimation is a random on-the-job search partial equilibrium model. In the model, firms post utility offers, consisting of a package of both amenities and pay. Workers can decide whether to accept these offers, where the idea is that if a worker voluntarily moves from one firm to another that implies that they value the new firm more than the previous one. Thus, the approach uses revealed preferences stemming from workers' voluntary transitions between firms to recover firm values. These transitions are used to construct a network of worker mobility across firms. The estimation procedure exploits the frequency and direction of these flows to recover a latent utility value for each firm that is consistent with the observed mobility patterns. Intuitively, firms that consistently attract workers from other firms, while rarely losing workers to lower-ranked firms, receive higher estimated values.

In implementing the method, we define voluntary employer-to-employer transitions as moves without unemployment benefits or low-earnings spells between adjacent firm spells. We apply the same earnings restriction as in our main AKM sample. Specifically, for each year, we calculate median annual earnings among men aged 45 and set earnings observations below 20 percent of this median to missing. We also exclude firms with fewer than five movers. We use the same definition of firms as in our main analysis. For firms with fewer than 1,000 unique workers, we use the firm identifier, whereas for firms with at least 1,000 unique workers, we use the establishment identifier. We define sectors using two-digit industry codes and use municipality as the geographic unit corresponding to the state variable in the original code. Finally, because information on unemployment benefit receipt is required to distinguish voluntary employer-to-employer transitions, the estimation sample is restricted to the period 1993–2018.

Following Sorkin (2018), we then retrieve a value for each firm, $V_j = \omega(\psi_j + lna_j)$, that consists of both a pay component ($\omega\psi_j$) and a non-pay component (a_j), where ω is the (unknown) unit conversion from money to utility.

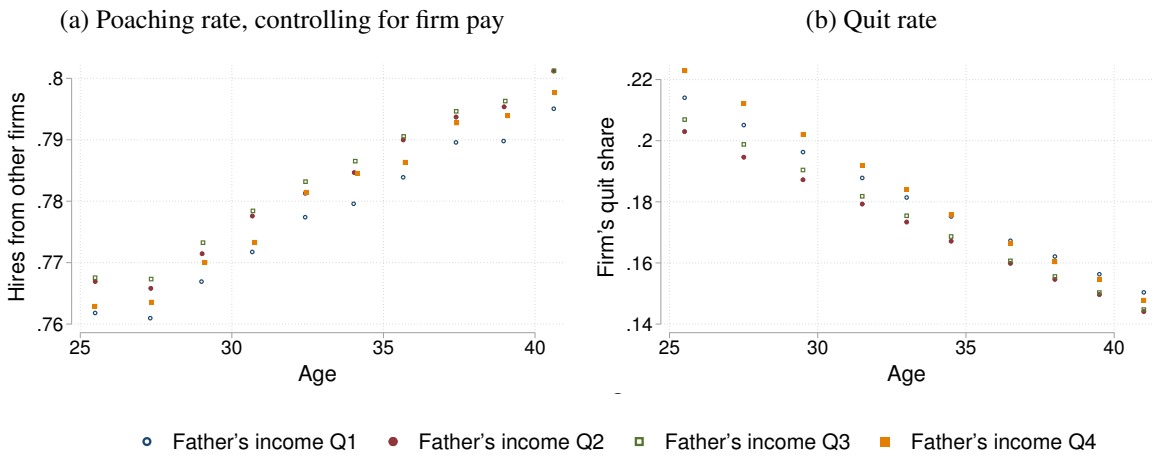
Table A14 provides results on how the estimated firm values are related to different worker characteristics. Men, workers with higher education, and workers with higher skill levels are more likely to work in firms with higher values. Figure A20 shows the mean firm fixed effect and firm values by industry. There is substantial heterogeneity in firm pay and firm values across industries. For example, the education industry exhibits relatively high

Table A14: Firm value and worker characteristics

	Firm value	Firm value	Firm value	Firm value
Woman	-0.026			
	(0.000)			
Years of education		0.013		
		(0.000)		
Cognitive ability			0.016	
			(0.000)	
Non-cognitive ability				0.014
				(0.000)
Observations	752,061	752,061	360,173	343,872

Notes: Regression results using the estimated firm value as the outcome and different worker characteristics as the independent variables. Robust standard errors in parentheses.

Figure A19: Poaching and quit rates over the lifecycle

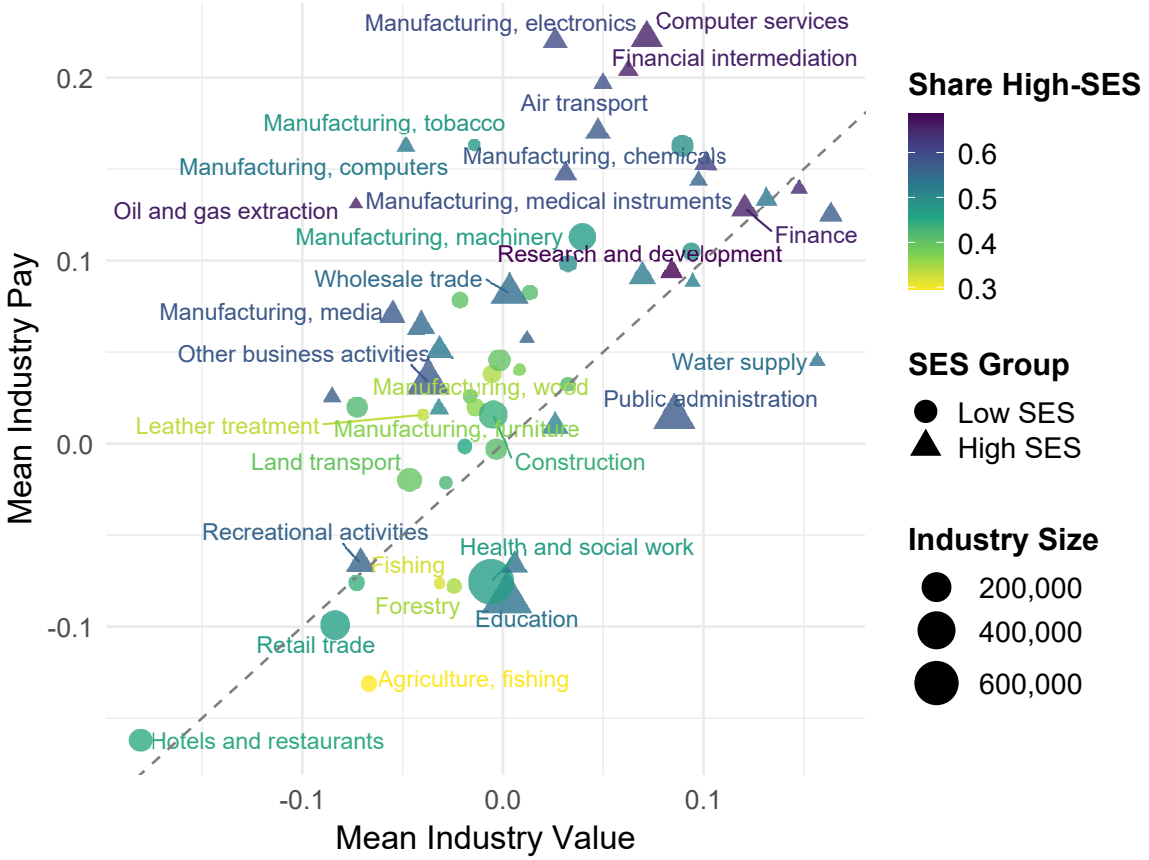


Notes: Subfigure (a) shows the mean share of hires from employment (poaching rate), controlling for the estimated firm premium $\hat{\psi}_j$. Subfigure (b) shows the quit rate at the firm.

firm values compared to its wage component, whereas oil and gas extraction exhibits relatively low firm values relative to pay, consistent with compensating differentials. In contrast, industries such as finance rank high in both wages and firm values, while the restaurant industry ranks low on both dimensions.

The firm values estimated following Sorkin (2018) are correlated with other measures of firm desirability, such as poaching rates from other firms and quit rates at the firm (Figure A19), and with the firm pay premium (Table A15).

Figure A20: Firm pay premia and firm values by industry



Notes: Mean firm pay and mean firm value in different industries, as calculated from the estimated AKM firm 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's income above the median.

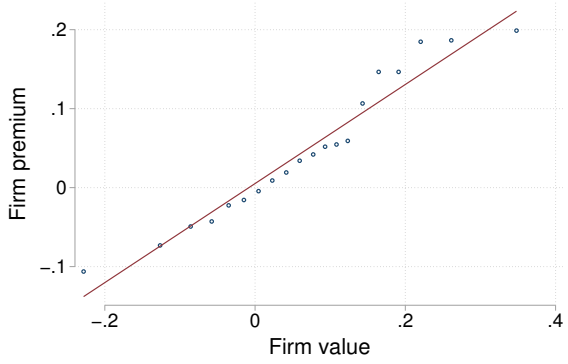
Table A15: Correlations between different measures of firm values

	$\hat{V}_j$	$\hat{\psi}_j$	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

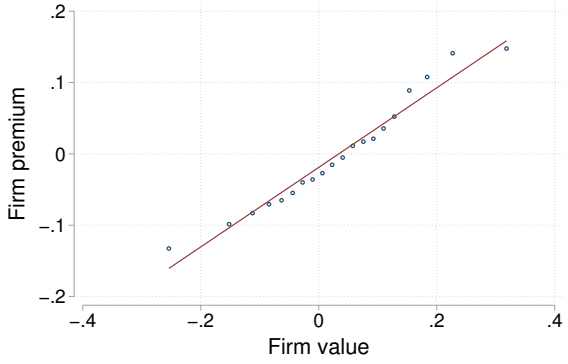
Notes: Correlations between estimated firm values $\hat{V}_j$ (estimated following Sorkin, 2018), firm fixed effects $\hat{\psi}_j$ (estimated from equation (1)), poaching rates (defined as the share of hires from employment), and quit rates at the firms.

Figure A21: Firm values and firm pay premia

(a) Father's income above the median



(b) Father's income below the median



Notes: Mean firm pay (estimated from equation (1)) against binned firm values (estimated following Sorkin, 2018). Subfigure (a) shows the results for children whose father's income is above the median, while subfigure (b) shows results for children with father's income below the median.

Table A16: 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 following Sorkin, 2018), by father's income quartile.

A5.2 Firm values and parental income

This appendix reports how the relationship between estimated firm values and firm pay varies by parental income. Figure A21 plots firm values against firm pay premia separately for children with fathers above and below the median income, and Table A16 reports the corresponding correlations by quartile of father's income. Table A17 repeats the analysis of SES gaps in firm values excluding small firms.

Table A17: SES gaps in firm values excluding small firms

	Dependent variable					
	$\hat{\psi}_{j=J(i,t)}$ (1)	$\hat{V}_j$ (2)	$\hat{V}_j$ (3)	$\hat{V}_j$ (4)	$\hat{V}_j$ (5)	$\hat{\psi}_{j=J(i,t)}$ (6)
$y_{f(i)}$	0.057	0.040	0.027	0.014	0.010	0.032
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\psi}_{j=J(i,t)}$				0.459	0.437	
				(0.001)	(0.001)	
$\hat{\alpha}_i$			0.112		0.044	
			(0.001)		(0.001)	
$\hat{V}_j$						0.607
						(0.001)
R^2	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 the 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 the father's log income. Columns (3)–(6) repeat these regressions with different sets of controls. Estimated firm values $\hat{V}_j$ are standardized to have mean zero and standard deviation one. Small firms are excluded in the table, where small firms are defined as the firms in the two bottom quintiles of numbers of employees. Robust standard errors in parentheses.

A6 Theoretical derivations

A6.1 Derivations for the model of Section 5.3

Throughout, all variables have mean zero, and u_i, v_i, w_i are mutually independent and independent of SES_i^*. We write $\sigma^2_{ses}, \sigma^2_u, \sigma^2_v, \sigma^2_w$ for the variances of SES_i^*, u_i, v_i, w_i, and recall $skill_i^* = \gamma_{ses} SES_i^* + u_i$.

Projection coefficients b_{skill} and b_{ses}

Firms project skill on the two signals $skill_i = skill_i^* + v_i$ and $SES_i = SES_i^* + w_i$. The required moments are

$$\begin{aligned}
\mathrm{Var}(skill_i) &= \gamma^2_{ses}\sigma^2_{ses} + \sigma^2_u + \sigma^2_v, & \mathrm{Cov}(skill_i, SES_i) &= \gamma_{ses}\sigma^2_{ses},\\
\mathrm{Var}(SES_i) &= \sigma^2_{ses} + \sigma^2_w, & \mathrm{Cov}(skill_i^*, SES_i) &= \gamma_{ses}\sigma^2_{ses},\\
\mathrm{Cov}(skill_i^*, skill_i) &= \gamma^2_{ses}\sigma^2_{ses} + \sigma^2_u.
\end{aligned}$$

The coefficients of this two-signal projection follow from the usual formula for a regression with two regressors:

$$b_{skill} = \frac{\mathrm{Cov}(skill^*, skill)\,\mathrm{Var}(SES) - \mathrm{Cov}(skill^*, SES)\,\mathrm{Cov}(skill, SES)}{\mathrm{Var}(skill)\,\mathrm{Var}(SES) - \mathrm{Cov}(skill, SES)^2},$$

and symmetrically for b_{ses}, interchanging *skill* and *SES* in the numerator. Substituting the moments above, the common denominator is

$$D = \gamma^2_{ses}\sigma^2_{ses}\sigma^2_w + \left(\sigma^2_u + \sigma^2_v\right)\left(\sigma^2_{ses} + \sigma^2_w\right), \tag{A2}$$

and the coefficients are

$$b_{skill} = \frac{\gamma^2_{ses}\sigma^2_{ses}\sigma^2_w + \sigma^2_u\left(\sigma^2_{ses} + \sigma^2_w\right)}{D}, \qquad b_{ses} = \frac{\gamma_{ses}\,\sigma^2_{ses}\,\sigma^2_v}{D}. \tag{A3}$$

Reliability-ratio form. Define

$$\sigma^2_{skill^*} = \gamma^2_{ses}\sigma^2_{ses} + \sigma^2_u, \quad \rho_{skill} = \frac{\sigma^2_{skill^*}}{\sigma^2_{skill^*} + \sigma^2_v}, \quad \rho_{ses} = \frac{\sigma^2_{ses}}{\sigma^2_{ses} + \sigma^2_w}, \quad \pi = \frac{\gamma^2_{ses}\sigma^2_{ses}}{\sigma^2_{skill^*}}. \tag{A4}$$

Using $\sigma^2_u = (1-\pi)\sigma^2_{skill^*}$, $\sigma^2_v = \sigma^2_{skill^*}(1-\rho_{skill})/\rho_{skill}$, and $\sigma^2_w = \sigma^2_{ses}(1-\rho_{ses})/\rho_{ses}$, the determinant (A2) becomes

$$D = \frac{\sigma^2_{skill^*}\,\sigma^2_{ses}}{\rho_{skill}\,\rho_{ses}}\left(1 - \pi\,\rho_{skill}\,\rho_{ses}\right), \tag{A5}$$

and substituting into (A3) yields the expressions used in the main text,

$$b_{skill} = \frac{\rho_{skill}\big(1 - \pi\, \rho_{ses}\big)}{1 - \pi\, \rho_{skill}\, \rho_{ses}}, \qquad b_{ses} = \gamma_{ses}\, \frac{\rho_{ses}\big(1 - \rho_{skill}\big)}{1 - \pi\, \rho_{skill}\, \rho_{ses}}. \tag{A6}$$

As a check, $\sigma_w^2 = 0$ ($\rho_{ses} = 1$) reduces (A6) to $b_{skill} = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ and $b_{ses} = \gamma_{ses}\sigma_v^2/(\sigma_u^2 + \sigma_v^2)$; $\rho_{skill} \to 1$ gives $(b_{skill}, b_{ses}) \to (1, 0)$; and $\rho_{skill} \to 0$ gives $b_{ses} \to \gamma_{ses}\rho_{ses}$.

Total and direct gradients and their lifecycle behavior

Throughout this subsection we omit the network term $\lambda\, SES_i^*$ of eq. (15), which adds λ to both the total and the direct gradient and drops out of all derivatives. With $\psi_{j=J(i,t)} = \delta\big(b_{skill} skill_i + b_{ses} SES_i\big)$ and $\mathrm{Cov}(skill_i, SES_i^*)/\mathrm{Var}(SES_i^*) = \gamma_{ses}$, as well as $\mathrm{Cov}(SES_i, SES_i^*)/\mathrm{Var}(SES_i^*) = 1$, the regression of ψ on SES_i^* has slope $\beta_{firm} = \delta\big(b_{skill}\gamma_{ses} + b_{ses}\big)$, while conditioning on $skill_i^*$ leaves $\beta_{firm,direct} = \delta\, b_{ses}$. Substituting (A6),

$$\beta_{firm} = \delta\gamma_{ses}\, \frac{\rho_{skill} + \rho_{ses} - (1+\pi)\rho_{skill}\rho_{ses}}{1 - \pi\rho_{skill}\rho_{ses}}, \qquad \beta_{firm,direct} = \delta\gamma_{ses}\, \frac{\rho_{ses}\big(1 - \rho_{skill}\big)}{1 - \pi\rho_{skill}\rho_{ses}}. \tag{A7}$$

Differentiating in ρ_{skill} (employer learning),

$$\frac{\partial\beta_{firm}}{\partial\rho_{skill}} = \delta\gamma_{ses}\, \frac{(1-\rho_{ses})(1-\pi\rho_{ses})}{\big(1 - \pi\rho_{skill}\rho_{ses}\big)^2}, \qquad \frac{\partial\beta_{firm,direct}}{\partial\rho_{skill}} = -\,\delta\gamma_{ses}\, \frac{\rho_{ses}\big(1 - \pi\rho_{ses}\big)}{\big(1 - \pi\rho_{skill}\rho_{ses}\big)^2}. \tag{A8}$$

For the first derivative, expanding the quotient-rule numerator cancels every term in ρ_{skill} and leaves $1-(1+\pi)\rho_{ses}+\pi\rho_{ses}^2 = (1-\rho_{ses})(1-\pi\rho_{ses})$. Since $\delta, \gamma_{ses} > 0$ and $\pi\rho_{ses} < 1$, the sign of $\partial\beta_{firm}/\partial\rho_{skill}$ is that of $(1 - \rho_{ses}) \geq 0$: the gradient rises with employer learning iff family background is observed imperfectly ($\rho_{ses} < 1$), and is flat at $\beta_{firm} = \delta\gamma_{ses}$ when $\rho_{ses} = 1$. The direct gradient $\beta_{firm,direct} \propto (1 - \rho_{skill})$ is strictly decreasing in ρ_{skill}.

The estimated direct effect when the researcher observes skill imperfectly

The decomposition of Section 5.1 conditions on *measured* skill rather than true skill. Let the firm premium satisfy $\psi_i = \beta_{firm,direct}\, SES_i^* + \theta\, skill_i^* + \eta_i$ with $\eta_i \perp (SES_i^*, skill_i^*)$, where $\theta = \delta b_{skill}$ is the skill loading. Suppose we observe SES_i^* but only a noisy skill proxy $\widetilde{skill}_i = skill_i^* + e_i$, with e_i as a white-noise error with $\mathrm{Var}(e_i) = \sigma_e^2$. Regressing ψ on $(SES^*, \widetilde{skill})$, the probability limit of the coefficient on SES^* is

$$\widehat{\beta}_{firm,direct} = \beta_{firm,direct} + \theta\, \gamma_{ses}\, \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2}. \tag{A9}$$

The bias is positive whenever the researcher's skill measure is noisy ($\sigma_e^2 > 0$): the proxy fails to absorb all of $skill_i^*$, and the residual skill sorting – which is correlated with background

through γ_{ses} – loads onto the estimated direct effect. It vanishes when skill is measured perfectly ($\sigma_e^2 = 0$) and rises to $\theta\gamma_{ses}$ (the full sorting term) as $\sigma_e^2 \to \infty$. Because the skill loading $\theta = \delta b_{skill}$ increases over the career (as ρ_{skill} rises), the bias term in (A9) grows with age, partially offsetting the decline in the true direct effect $\beta_{firm,direct}$ and helping reconcile the model with the near-constant empirical profile in Figure 7.

A6.2 Firm pay, firm values, and measurement error

An alternative interpretation of the conditional SES gradients in Table 5 is that our estimates of firm pay and firm value are affected by measurement error. Estimating these objects from job-to-job flows 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), but introduces an attenuation bias when they serve as controls on the right-hand side of a regression – a bias that spills over to the coefficient on parental income.

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

$$\hat{V}_j = V_j + u_j^V \tag{A10}$$

$$\hat{\psi}_j = \psi_j + u_j^\psi \tag{A11}$$

where u_j^V and u_j^ψ are classical measurement errors with variances $\sigma^2_{u^V}$ and $\sigma^2_{u^\psi}$, respectively. Assume further that firm value is the sum of firm pay and non-pay amenities,

$$V_j = \psi_j + a_j, \tag{A12}$$

and that firm pay increases in parental income,

$$\psi_j = \gamma y_f + \varepsilon_j \tag{A13}$$

with $\gamma > 0$ and ε_j capturing heterogeneity in firm pay orthogonal to y_f, with variance σ_ε^2.

Consider the regression of the estimated firm value $\hat{V}_j$ on the estimated pay premium $\hat{\psi}_j$ and parental income y_f – corresponding to column (4) of Table 5 – and suppose that amenities a_j are uncorrelated with parental income and firm pay, so that the true conditional gradient is zero. The coefficient β_y on parental income is nonetheless positive. Noting that $\text{Cov}(\hat{V}_j, y_f) = \text{Cov}(\hat{\psi}_j, y_f) = \gamma\sigma_y^2$, $\text{Cov}(\hat{V}_j, \hat{\psi}_j) = \text{Var}(\psi_j) = \gamma^2\sigma_y^2 + \sigma_\varepsilon^2$ and $\text{Var}(\hat{\psi}_j) =$

$\text{Var}(\psi_j) + \sigma^2_{u^\psi}$, the standard formula for a regression with two regressors yields

$$\beta_y = \frac{\text{Cov}(\hat{V}, y_f)\,\text{Var}(\hat{\psi}) - \text{Cov}(\hat{V}, \hat{\psi})\,\text{Cov}(\hat{\psi}, y_f)}{\text{Var}(y_f)\,\text{Var}(\hat{\psi}) - \text{Cov}(\hat{\psi}, y_f)^2} = \frac{\gamma\sigma^2_y\,\sigma^2_{u^\psi}}{\sigma^2_y\left(\sigma^2_\varepsilon + \sigma^2_{u^\psi}\right)} = \gamma\frac{\sigma^2_{u^\psi}}{\sigma^2_\varepsilon + \sigma^2_{u^\psi}}, \quad \text{(A14)}$$

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

Split-sample instrumental-variable approach

To assess the role of sampling error empirically, we implement a split-sample instrumental-variables (IV) approach, similar to 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.[50] Because we can only include firms observed in both subsamples, this procedure reduces the number of firms and individual observations relative to the baseline.

Panel C of Table 5 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 we instrument the estimated firm value in column (2) and the estimated firm pay in column (4), the coefficients on these controls increase and the coefficient on parental income declines correspondingly. This pattern is consistent with classical measurement error in the firm-level regressors, which the IV procedure helps correct. The magnitude of the adjustment is relatively small – around 10 percent – and the SES gradients remain positive and substantial.[51]

We therefore conclude that sampling noise is not a major driver of the conditional SES gradients in Table 5. 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.

[50] We estimate firm values following Sorkin (2018) independently in each subsample, obtaining two measures of firm values, $\hat{V}_1$ and $\hat{V}_2$. We also estimate separate AKM models in the two subsamples to obtain corresponding firm premia, $\hat{\psi}_1$ and $\hat{\psi}_2$. We then use $\hat{V}_1$ as an instrument for $\hat{V}_2$, and $\hat{\psi}_1$ as an instrument for $\hat{\psi}_2$. Using the firm values as instruments yields a first stage of 0.888; using the firm premia, a first stage of 0.921.

[51] The results are also robust to excluding small firms, for which the estimated firm values are more likely to be noisy (Table A17).