What Drives Wage Sorting? Evidence From West Germany*

Andre Mouton
Wake Forest University

February 2025 (Current Version Here)

Abstract

A key source of rising inequality is *wage sorting*: in many countries, high-earning workers are increasingly sorted into high-paying firms. This paper combines German survey and administrative data to study the drivers of wage sorting. I find that it is primarily an industry-level phenomenon, with evidence rejecting an assortative matching mechanism. Wage sorting strengthened over 1993-2017 as two major labor market trends - rising skill premia and falling manufacturing employment - interacted with sectoral pay gaps. I show that wage sorting is predicted by human capital and firm investment, which are strongly correlated across industries. These relationships are rationalized through a simple rent-sharing model, with evidence supporting a pair of general mechanisms.

Keywords: Wage Inequality, Firm-Wage Differentials, Labor Sorting
JEL Codes: E24, J21, J24, J31

*Email: moutona@wfu.edu. Previously circulated as "What Drives West German Wage Sorting?" This paper benefited greatly from the advice of Ali Shourideh, Laurence Ales, Rebecca Lessem, and Brian Kovak, and helpful comments by Aeimit Lakdawala, Emily Moschini, and seminar participants at Carnegie Mellon, Wake Forest, the Liberal Arts Macroeconomics Conference, and SOLE. I'm grateful to the Research Data Center of the Federal Employment Agency at the Institute for Employment Research (IAB-FDZ) for their generous assistance and provision of data, which was made available under contract number FDZ1747.

1 Introduction

Two stylized facts emerge from recent studies of the wage distribution. First, much observed dispersion in wages is due, not to differences between workers, but to differences between firms. Studies of matched employee-employer data find that some firms pay consistently higher wages to their workers, relative to what those workers earn at past or future employers.[1] Second, in many countries high-earning individuals are more likely to work for high-paying firms. This phenomenon, known as *wage sorting* (Bagger, Sørensen and Vejlin, 2013), is a source of rising wage inequality in the United States and Europe.[2]

Despite its importance, we know little about wage sorting. It is seemingly at odds with conventional stories about post-1980's wage inequality, which focus on demand-side forces related to technological and structural change. In that literature, rising inequality is equated with an increase in skill premia, with no role for differences in what firms pay. At the same time, wage sorting is challenging to study. Matched datasets are typically administrative in origin, and lack detailed information on firm and worker characteristics. The canonical empirical framework - after Abowd, Kramarz and Margolis (1999) - also suffers from issues related to statistical bias, and is known to be inconsistent with search-theoretic models of labor sorting. For these reasons, connecting theory and data is difficult.

This paper combines West German survey and administrative data to shed light on the mechanisms driving wage sorting, and on the reasons for its growing importance. West Germany is a natural area of focus. In a seminal work, Card, Heining and Kline (2013) showed that wage sorting accounts for one-third of the rise in German wage variance since the 1990's, a finding that I replicate in Figure 1.1. Similar results have been established for the U.S. and elsewhere, but German data offer a unique combination of accessibility, quality, and coverage. I use this data to answer three questions. First, is wage sorting a match-level phenomenon consistent with the theoretical search literature, or is it associated with occupational, sectoral, and/or geographic sorting? Second, is wage sorting related to, or distinct from, conventional demand-side explanations for rising wage inequality? And third, what are the observable aspects of worker-firm sorting, and with what underlying mechanisms are they consistent?

The methodological approach I take is forensic in that, given our limited knowledge, I consider a range of theoretically-motivated hypotheses that I evaluate through reduced-form methods. The goal of the analysis is to identify the mechanism(s) associated with wage sorting, defined as the covariance of the person and employer wage effects from a standard

[1] See Card, Cardoso, Heining and Kline (2018) for a survey.

[2] See discussion below. Closely related is a rise in between-firm wage inequality; see Dunne, Foster, Haltiwanger and Troske (2004), Simón (2010), Barth, Bryson, Davis and Freeman (2016), and Tomaskovic-Devey et al. (2020).

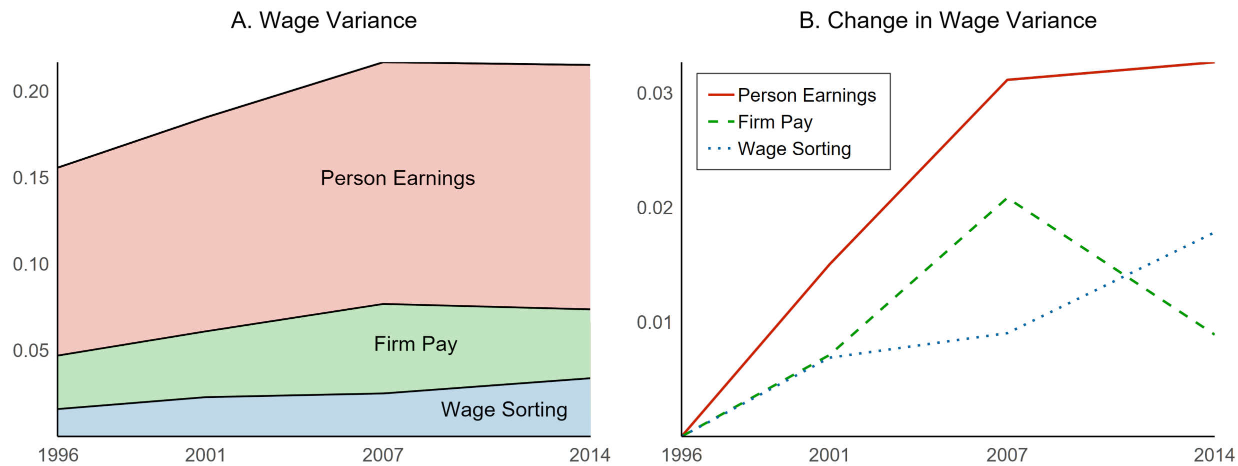

Figure 1.1: West German Wage Variance Components, 1993-2017

DATA: German Linked Employer-Employee Dataset (LIAB) and IAB-estimated AKM effects. NOTE: AKM variance components (Abowd et al., 1999). Wage effects estimated in four panels (1993-99, 1998-04, 2003-10, 2010-17) with variance components plotted over panel mid-point. Residual and time-varying effects omitted. See Section 3 for details.

AKM regression (Abowd et al., 1999). This approach limits the paper's scope in several respects. The first is that I cannot and do not attempt to rule out any mechanism *in general*, but only as an explanation for the outcome of interest. Secondly, while measures of wage sorting are known to be biased downwards and therefore to give an incomplete picture (Andrews, Gill, Schank and Upward, 2008), this paper is nevertheless concerned with the patterns seen in previous studies and shown in Figure 1.1. I rule out bias as an explanation for the observed trend, and I establish the bias-robustness of the main results, but *unobserved* wage sorting is not my focus.

To implement this approach, I begin with a conceptual framework that bridges the gap between theory and data. Workers and firms split match surplus, which varies across matches due to market-level entry costs (*e.g.* education or fixed capital), idiosyncratic productivity and ability, and complementarities in match production. The framework yields closed-form analogues of the AKM wage effects, allowing me to characterize the relationships between wage sorting and several broad classes of worker-firm sorting: *assortative matching* of high-ability workers and high-productivity firms as in Shimer and Smith (2000), *technical sorting* into markets in which both workers and firms face high or low entry costs, and *sorting on unobservables* in which high-ability workers or high-productivity firms sort into high-cost markets. Critically, the three mechanisms give different predictions about the within- versus between-market components of wage sorting.

The paper's first main result is that wage sorting occurs across sectors and - to a lesser extent - occupations, and is technical rather than assortative in nature, ruling out the most obvious set of theoretical explanations.[3] Group covariance decompositions show wage sorting to be entirely between aggregated industry and occupation groups. It is absent within industry-occupation cells, where the conceptual framework predicts that assortative matching is most relevant. I establish the robustness of the group-level results to "limited-mobility bias" (Andrews et al., 2008), and show that group-mean wage effects are strongly associated with observable (explicit and implicit) entry costs, consistent with a technical mechanism. On the other hand, wage sorting appears to be unimpeded by the type of search frictions assumed in the assortative matching literature: it is at least as prevalent among new matches, entering establishments, and first-time workers.

Second, I find that the increase in wage sorting over 1993-2017 is almost entirely accounted for by developments attributable to industry and occupation demand, suggesting that conventional explanations for rising wage inequality are correct, but incomplete. Trend decompositions show half of the increase to be the result of growing service sector employment. Much of this has occurred within temporary employment agencies, which supply labor predominantly to manufacturing establishments; hence domestic outsourcing is likely to play a role. Nearly as important as sectoral composition is a rise in worker earnings within high-paying, high-skill sectors. A modified Oaxaca-Blinder analysis indicates that this is not due to widening skill differentials, but to an increase in the return to education. Both results are consistent with past work linking occupation and industry demand shifts to German wage inequality (Dauth, Findeisen and Suedekum, 2017; Dustmann, Ludsteck and Schönberg, 2009),[4] but they indicate a central role for sectoral gaps in firm pay, which are absent from canonical models of skill-bias.

Third, I characterize the market features associated with wage sorting, and provide evidence on a pair of causal mechanisms that, if present, are likely to extend beyond West Germany. The conceptual framework attributes technical sorting to greater entry (*i.e.* due to supply or demand) in markets where upfront costs are high or low for both workers and firms. I find that capital - physical and human - predicts wage sorting along two margins. First, IT-intensive industries are also education-intensive. This is not due to more managerial or less routine labor, but to a greater share of research-related occupations and tasks, consistent with a mechanism in which knowledge production requires bilateral capital investments. Second, capital-intensive (and high-paying) industries have lower match separation rates, resulting in an older and more experienced workforce. Age and match tenure are both associated with growth in worker earnings, suggesting that higher pay raises the level of

[3] See Section 5B of Song et al. (2018) for a discussion of relevant assortative matching mechanisms.

[4] See also Spitz-Oener (2006); Bachmann and Burda (2010).

informal human capital in a sector by increasing retention.

Related literature. This paper contributes to an empirical literature that documents wage sorting's contribution to inequality, but provides little evidence on its provenance. Card et al. (2013) showed that wage sorting among West German men rose from a negligible level in the 1980's to one-sixth of wage variance by the early 2010's. Similar results have been shown for the United States (Song, Price, Guvenen, Bloom and von Wachter, 2018), Denmark (Bagger et al., 2013), and Sweden (see Håkanson, Lindqvist and Vlachos, 2021, app. D). Wage sorting is positive but declining in several countries where wage inequality is stable or falling (Torres, Portugal, Addison and Guimarães, 2018; Alvarez, Benguria, Engbom and Moser, 2018). Card et al. (2013) speculate that a decline in collective bargaining coverage may have contributed to German wage trends, and Song et al. (2018) suggest a variety of assortative matching mechanisms. Instead, I find that German wage sorting reflects stable industry pay and skill gaps, that have grown in importance because of changes to sectoral composition and a general rise in skill premia.

The finding of strong interactions between labor demand and firm-wage differentials has substantive implications for the inequality literature, in which these are usually studied separately. Canonical models of skill-bias assume perfectly competitive markets, with the effect of a shift in relative demand being solely to raise or lower the price of skill (Krusell, Ohanian, Ríos-Rull and Violante, 2000; Acemoglu and Autor, 2011; Autor and Dorn, 2013; Acemoglu and Restrepo, 2018). Conversely, market segmentation and relative demand are generally absent from studies relating wage dispersion to market institutions (Fortin and Lemieux, 1997; Lise, Meghir and Robin, 2016; Engbom and Moser, 2022) and worker-firm sorting (Hagedorn, Law and Manovskii, 2017; Bagger and Lentz, 2019; Bonhomme, Lamadon and Manresa, 2019). This paper's results suggest that the first literature misses an important channel by which demand affects wages. The implications for the second is less clear, as these papers are generally concerned with firm-pay gaps arising from market institutions or match effects. I find these to be unrelated to wage sorting, which is associated instead with differences in capital investment.

That wage sorting is mostly sectoral in nature is surprising, given the literature's focus on skill premia and occupational wage gaps; however it is consistent with a variety of empirical studies. Industries have long been known to vary both in terms of firm pay (Krueger and Summers, 1988; Gibbons and Katz, 1992; Abowd, Kramarz, Lengermann, McKinney and Roux, 2012) and workforce skill (Haskel and Slaughter, 2002; Buera, Kaboski, Rogerson and Vizcaino, 2022). This paper establishes that in Germany, the two gaps are correlated. Firm-wage differentials are traditionally attributed to differences in capital intensity and scale

(Krueger and Summers, 1988; Brown and Medoff, 1989), as well as institutional features such as collective bargaining agreements. I find that capital intensity, in particular, is strongly associated with German wage sorting. Finally, this paper's finding that skill and pay gaps are associated with IT investment, but also stable over time, is consistent with well-known results from Doms, Dunne and Troske (1997).

This paper is most directly related to recent work by Haltiwanger, Hyatt and Spletzer (2024), who find that much of the rise in U.S. wage sorting is also between-industry - an outcome suggesting that this paper's results are likely to generalize beyond Germany. While those authors are concerned with characterizing changes to between- and within-industry wage dispersion, and this paper with uncovering the mechanisms that drive wage sorting, we document a similar development: a shift in labor market composition towards high-paying, high-skill, and high-tech industries, and towards low-paying, low-skill service industries. The similarity of the German and U.S. experiences is consistent with this paper's contention that wage sorting is not an idiosyncratic outcome, but can be understood in terms of a pair of general mechanisms.

I begin by developing the conceptual framework in Section 2, followed by an overview of the datasets and methodological approach in Section 3. In Section 4 I evaluate the evidence for alternative wage sorting mechanisms, and in Section 5 I decompose the sources of rising wage sorting over 1993-2017. A set of specific hypotheses are evaluated in Section 6. Proofs of the main results are given in Appendix A, and auxiliary results in the Online Appendix.

2 Conceptual Framework

Here I consider the potential mechanisms by which wage sorting might arise, in the context of a simple one-period model where (1) rents are shared between workers and firms, (2) there is free entry at the market level, (3) entry costs vary across markets, and (4) firm productivity and worker ability are heterogeneous within, and potentially between markets. I take equilibrium worker distributions as given, and consider which types of worker-firm sorting patterns are capable of generating wage sorting. To this end I define wage effects analogous to those in the AKM literature, and give results on three candidate mechanisms.

Firms are assumed to operate in two differentiated *output markets* $z \in \{z_L, z_H\}$, and to draw an idiosyncratic productivity term $\zeta \in \{\zeta_-, \zeta_+\}$ upon creation of a new vacancy. Workers choose between two segmented *labor markets* $s \in \{s_L, s_H\}$ and draw ability $\sigma \in \{\sigma_-, \sigma_+\}$. I make several simplifying assumptions. First, I assume that ζ and σ are drawn subsequent to entry, however I allow the joint distribution $f_{z,s}(\zeta, \sigma)$ to vary arbitrarily across markets, taking matching as given.[5] Second, I abstract from market-clearing and take equilibrium employment shares $g_{z,s}$ as also given. Third, I assume that vacancy creation requires payment of a fixed cost $K(z)C(s)$ that is separable in types and paid by the firm,[6] where $K(z_H) > K(z_L)$ and $C(s_H) > C(s_L)$. Lastly, I do not model wage formation but instead assume that workers receive a share ω of match revenue, as would arise *e.g.* in a setting with wage bargaining subsequent to entry.

Free entry provides structure on wages, which take the form[7]:

$$w(z,s,\zeta,\sigma) = \frac{\omega}{1-\omega}\hat{y}(z,s,\zeta,\sigma)K(z)C(s) \ ,$$

where $\hat{y}(z,s,\zeta,\sigma) = \frac{y(\zeta,\sigma)}{\sum_{\zeta,\sigma} f_{z,s}(\zeta,\sigma)y(\zeta,\sigma)}$. Market-level (*i.e.* $z \times s$) productivity differentials are priced out, and therefore wage depends only on the worker's rent share, relative match productivity $\hat{y}$, and entry costs. I define the *person wage effect* (PE) and *firm wage effect* (FE) analogously to their AKM counterparts, as the mean log wage differentials associated with (s,σ) and (z,ζ) respectively:

$$PE(j,\sigma) \equiv \sum_{z',s',\zeta',\sigma'} g(z',s')f_{z',s'}(\zeta',\sigma')\Big[\log w(z',s,\zeta',\sigma) - \log w(z',s',\zeta',\sigma')\Big]$$
$$FE(i,\zeta) \equiv \sum_{z',s',\zeta',\sigma'} g(z',s')f_{z',s'}(\zeta',\sigma')\Big[\log w(z,s',\zeta,\sigma') - \log w(z',s',\zeta',\sigma')\Big] \ .$$

By construction, w will be log additive in PE, FE, and a mean-zero match term:

Proposition 1 (Wage Function). *The log wage function can be written, up to a normalization, as*

$$\log w(z,s,\zeta,\sigma) = \underbrace{PE(s,\sigma)}_{\textit{Person Effect}} + \underbrace{FE(z,\zeta)}_{\textit{Firm Effect}} + \underbrace{ME(z,s,\zeta,\sigma)}_{\textit{Match Effect}} \ ,$$

where

$$PE(s,\sigma) = \log C(s) + \sum_{z,\zeta}\left[\sum_{s',\sigma'} g(z,s')f_{z,s'}(\zeta,\sigma')\right]\log \hat{y}(z,s,\zeta,\sigma) \tag{1}$$

[5]Note that while f allows for a flexible distribution of match types, the assumption of free entry will continue to result in a zero-expected profits condition. If instead firms observe ζ prior to entry, then a zero-profit condition will hold only for the marginal entrant. Nevertheless the main implication of free entry - that higher mean values of ζ are offset by lower output prices - would continue to hold in this setting, at last partially.

[6]Separability is convenient for analytic results, the more important restriction being strict monotonicity in both arguments. That $C(s)$ is paid by the firm is again convenient, but largely WLOG given that the model is static.

[7]Derivations are given in Appendix C.

$$FE(z,\zeta) = \log K(z) + \sum_{s,\sigma} \left[\sum_{z',\zeta'} g(z',s) f_{z',s}(\zeta',\sigma) \right] \log \hat{y}(z,s,\zeta,\sigma) \tag{2}$$

$$\begin{aligned} ME(z,s,\zeta,\sigma) = \log \hat{y}(z,s,\zeta,\sigma) &- \sum_{s,\sigma} \left[\sum_{z',\zeta'} g(z',s) f_{z',s}(\zeta',\sigma) \right] \log \hat{y}(z,s,\zeta,\sigma) \\ &- \sum_{z,\zeta} \left[\sum_{s',\sigma'} g(z,s') f_{z,s'}(\zeta,\sigma') \right] \log \hat{y}(z,s,\zeta,\sigma) \ . \end{aligned} \tag{3}$$

Proof. See Appendix A.1.

Note that while Proposition 1 maps model primitives into a set of log additive wage effects, this is done by construction, and it does not imply any consistency with the assumptions of the benchmark AKM model. This can be seen in the fact that PE and FE depend on match- and market-specific terms. Critically, the AKM assumption of wage separability will only hold whenever $ME(z,s,\zeta,\sigma) \equiv 0$. If match effects are instead present - for example, if ζ and σ are complements in production - then $\hat{y}$ will be apportioned between PE and FE in an opaque manner. Additionally, as y is unrestricted, it is not assumed that FE is monotonic with respect to ζ, or PE with respect to σ. Hence the framework can capture in rough fashion the non-monotonicities discussed in de Melo (2018).

A positive covariance of the wage effects PE and FE may arise from any of several worker-firm sorting patterns, shown in Figure 2.1. I refer to these as *wage sorting mechanisms*, and in Appendix C.2 I show that sufficient conditions are as follows:

1. **Technical sorting**. High-K firms employ more high-C workers: $\frac{g(z_H,s_H)}{g(z_H,s_L)} > \frac{g(z_L,s_H)}{g(z_L,s_L)}$.

2. **Positive assortative matching**. High-ζ firms employ more high-σ workers: $\frac{\partial y(\zeta,\sigma)}{\partial \zeta} > 0$ and $\frac{\partial y(\zeta,\sigma)}{\partial \sigma} > 0$ for all ζ and σ, and $\frac{f_{z,s}(\zeta_+,\sigma_+)}{f_{z,s}(\zeta_+,\sigma_-)} > \frac{f_{z,s}(\zeta_-,\sigma_+)}{f_{z,s}(\zeta_-,\sigma_-)}$ for all z and s.

3a. **Sorting on unobservables - ability**. High-K firms employ more high-σ workers: $K(z_H) - K(z_L)$ is sufficiently large, $\frac{\partial y(\zeta,\sigma)}{\partial \sigma} > 0$ for all ζ and σ, and for any s, s', ζ, and ζ' we have $\frac{f_{z_H,s}(\zeta,\sigma_+)}{f_{z_H,s}(\zeta,\sigma_-)} > \frac{f_{z_L,s'}(\zeta',\sigma_+)}{f_{z_L,s'}(\zeta',\sigma_-)}$.

3b. **Sorting on unobservables - productivity**. High-C workers work at high-ζ firms: $C(s_H) - C(s_L)$ is sufficiently large, $\frac{\partial y(\zeta,\sigma)}{\partial \zeta} > 0$ for all ζ and σ, and for any z, z', σ, and σ' we have $\frac{f_{z,s_H}(\zeta_+,\sigma)}{f_{z,s_H}(\zeta_-,\sigma)} > \frac{f_{z',s_L}(\zeta_+,\sigma')}{f_{z',s_L}(\zeta_-,\sigma')}$.

Technical sorting reflects greater entry in markets (*e.g.* due to relative demand) where costs C and K are either both high or both low. *Positive assortative matching* requires $y(\zeta,\sigma)$ monotonically increasing, and sorting of high-σ workers into high-ζ firms within-market. If sorting instead occurs across markets, then higher values of y will induce entry and lower values of FE, and wage sorting will not occur.[8] *Sorting on unobservables* represents an intermediate case, in which idiosyncratic firm or worker heterogeneity is associated with higher fixed costs on the other side of the market. For the reason just described, this mechanism will tend to generate *negative* wage sorting unless K- or C-differentials are sufficiently large.

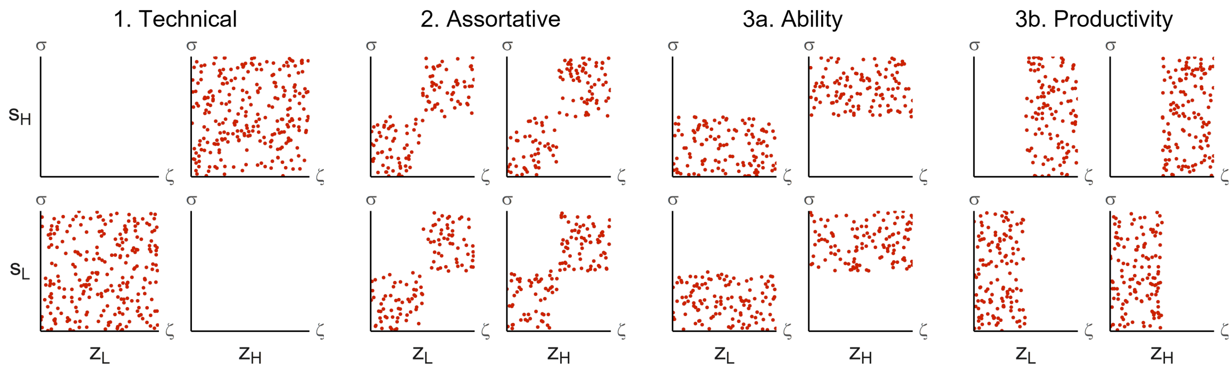

Figure 2.1: Wage Sorting Mechanisms

The model in turn yields a set of predictions regarding the within- versus between-market components of each wage sorting mechanism:

Proposition 2 (Wage Sorting Mechanisms). *Under free entry, the covariance of PE and FE will satisfy the following properties:*

- Technical sorting*: between-z, between-s, and absent within-(z, s).*
- Positive assortative matching*: within-(z, s) and absent between-z and between-s.*
- Sorting on unobservables*: between-s or between-z, and absent within-(z, s).*

Proof. See Appendix A.2.

The covariance of PE and FE will be both between-z and between-s in the case of technical sorting, entirely within-(z, s) for positive assortative matching, and either between-z or between-s in the remaining two cases.

The conceptual framework has two useful empirical implications. First, if one has data that captures market segmentation (z and s), then group covariance decompositions will provide a means of differentiating between the candidate mechanisms. Second, the model predicts that between-market wage sorting requires differentials in costs (K and/or C), which

[8]Free entry is a strong assumption, which will fail to hold when, for example, ζ is known prior to entry. Therefore these statements should be interpreted as tendencies, which will hold to the extent that the assumption of free entry is correct. See Appendix C for additional discussion. The strength of this assumption is one reason I consider a broad set of hypothesis tests when evaluating wage sorting mechanisms in Section 4.

implies a relationship in terms of observable characteristics that can, in principle, be studied directly. Hence while the model described above is reduced-form and somewhat stylized, it provides a degree of traction for the empirical analyses that follow.

3 Data, Methodology, and Trend

This section describes the data and the AKM methodology, and characterizes the trend of interest. I begin with an overview of the datasets used in the paper. This is followed by a brief description of the AKM approach, and a replication of the results in Card et al. (2013). I close with results on the incidence of limited-mobility bias, and a discussion of the paper's approach in addressing this issue.

3.1 Data

LIAB Linked Employer-Employee Dataset. The primary dataset used is the German linked employee-employer dataset (LIAB), provided courtesy of the Institute for Employment Research (IAB).[9] The IAB conducts an annual survey of German business establishments, collecting information on operational activities, investment, and hiring.[10] The IAB survey is then linked to administrative (and primarily social security) records for all workers formally employed at surveyed establishments, which include the person's occupation, wage earnings, and basic information on demographics and educational attainment. In a given year, the LIAB contains between 4 and 15 thousand establishments, and between 1.5 and 2.5 million employed persons, representing approximately 5% of the German workforce.

Following Card et al., I restrict the sample to full-time West German workers aged 20-60, and I impute top-coded wages through a set of Tobit regressions.[11] While establishment non-response rates are generally low, they vary considerably across variables and years within the IAB survey. Therefore to minimize issues related to selection, I impute missing responses using a regression approach with controls for year, industry, and establishment size.[12]

[9] LIAB cross-sectional model 2, version 1993-2017 (Schmidtlein et al., 2019). DOI: 10.5164/IAB.LIABQM29317.de.en.v1.

[10] An establishment is defined as a physical workplace, though locations may be aggregated when they share the same corporate ownership, industry classification, and municipal code. The sampling design is stratified, and therefore all results in this paper make use of survey weights.

[11] The sample restrictions allow for greater comparability over time, as discussed in Appendix B.1. Wages are top-coded at the annual social security contribution thresholds, which vary by year. The imputations performed for this paper appear only in Table 3.1 and a small set of auxiliary results, however an identical procedure is used by IAB prior to calculating the AKM wage effects discussed below. Note that the IAB-provided AKM effects (below) impose identical sample restrictions, save for the inclusion of apprenticeships.

[12] I perform OLS, fractional logit, or logit regressions depending upon the variable in question. Imputations primarily affect the results in the Appendix, as the variables analyzed in Section 6 have high response rates.

IAB-estimated AKM Wage Effects, Linked to the LIAB. Also provided by the IAB are an updated set of AKM person and firm effects from Card et al..[13] These are estimated on a larger dataset consisting of the universe of workers subject to social security contributions, and then matched with the LIAB subsample. Below I refer to the combined dataset as the "LIAB-AKM". Although the AKM wage regression can be implemented directly on the LIAB, use of the provided wage effects is preferable for reasons that I discuss below.

To satisfy disclosure requirements and to minimize the impact of coding changes over time,[14] I construct aggregate industry and occupation codes that preserve as much as possible of the 3-digit level variation in the AKM wage effects. To this end I combine neighboring 3-digit codes that exhibit similar mean person and employer effects. A set of 12 industry and 15 occupational groups is sufficient to capture most (nine-tenths) of the between-group variances of the wage effects, highlighting the aggregate nature of the wage structure characterized in Section 4.[15] Appendix B.1 contains details on the aggregation process.

2006 BIBB/BAuA Employment Survey. In Section 6, establishment-level data are supplemented with task and human capital data obtained from a 2005-2006 survey, conducted by the Federal Institute for Vocational Education and Training (BIBB) in partnership with the Federal Institute for Occupational Health and Safety (BAuA).[16] The survey draws on a random sample of the employed German labor force, asking respondents a range of questions concerning - among other things - job tasks and human capital requirements. Response categories are either binary or in the form of verbal frequencies. In the latter case I assign numerical values, and I then aggregate the BIBB survey to the industry-occupation level through regressions on industry and occupation fixed effects, plus a set of sector-occupation interactions.[17] For consistency I apply the same sample selection criteria as with the LIAB, though differences in survey design prevent absolute comparability between the two samples.

[13] Bellman et al. (2020).

[14] Industry codes are broadly comparable over time, however a redesigned occupational coding scheme is employed after 2010. Time-consistent codes are provided by IAB, however these rely on imputation. To reduce inconsistencies, I propagate industry and occupation codes forwards or backwards (as applicable) whenever a job spell is observed on both sides of a coding change.

[15] The constructed industry groups are similar to NACE sections, but with greater differentiation among manufacturing sectors. A notable difference is that, because few temporary agency establishments are observed in the 1990's, I merge this industry group with hospitality and accommodation. Occupational groups lack a natural comparison as the German coding system is not hierarchical. A previous set of aggregated codes was proposed by Blossfeld (1985), however I find that these perform poorly at capturing the between-occupation variation in the firm AKM effect.

[16] BIBB/BAuA-Erwerbstätigenbefragung 2006, version 4820 (Hall and Tiemann, 2021). DOI: 10.4232/1.11072.

[17] Verbal frequencies consist of either 3 or 4 values (*e.g.* "rarely", "always"), which I assume to lie evenly spaced on the unit interval. Industry (12) and occupation (15) classifications are the same as described above, and interactions consist of 3 industry and 4 occupational groups. While fixed effects regressions reduce noise in industry-occupation cells with few observations, these cells account for only a small portion of jobs, and therefore similar results are obtained when using simple means.

3.2 German Wage Sorting

I next describe the AKM methodology used to measure wage sorting, and I establish that the upward trend documented by Card et al. is present in the LIAB over the period studied. The AKM wage decomposition, after Abowd et al. (1999), begins with a regression of the form

$$w_{p,t} = \pi_p + \phi_{f(p,t)} + x'_{p,t}\beta + \epsilon_{p,t} \ , \tag{4}$$

where $w(p,t)$ is the log daily wage of person p in year t, π and ϕ are time-invariant wage effects associated with person p and their time-t employer, and x is a vector containing year fixed-effects and a cubic polynomial in worker age, interacted with dummies for educational attainment as in Card et al.. Regression (4) is implemented by the IAB in four partially-overlapping panels, each spanning 7-8 years. Wage variance can then be decomposed as the sum of the variances and covariances of the estimated regression effects:

$$\begin{aligned} Var(w_{p,t}) = {} & Var(\pi_p) + Var(\phi_{f(p,t)}) + Var(x'_{p,t}\beta) + Var(\epsilon_{p,t}) \\ & + \underbrace{2Cov(\pi_p, \phi_{f(p,t)})}_{\text{Wage Sorting}} + 2Cov(\pi_p, x'_{p,t}\beta) + 2Cov(\phi_{f(p,t)}, x'_{p,t}\beta) \ . \end{aligned} \tag{5}$$

Taking first differences of equation (5) allows one to study the sources of rising wage variance. The results of this decomposition are shown in Table 3.1.

Table 3.1: AKM Variance Decomposition, 1993-2017

	1993-99	1998-04	2003-10	2010-17
$Var(w)$	.1684	.1997	.2321	.2316
$Var(\pi)$	.1088	.1238	.1399	.1415
$Var(\phi)$	.0310	.0381	.0518	.0399
$Var(x'\beta)$	.0039	.0054	.0054	.0131
$Var(\epsilon)$	.0126	.0149	.0157	.0181
$2\times Cov(\pi,\phi)$	.0159	.0228	.0249	.0337
$2\times Cov(\pi,x'\beta)$	.0018	-.0001	-.0001	-.0186
$2\times Cov(\phi,x'\beta)$	.0013	.0015	.0024	.0008
$Corr(\pi,\phi)$	.137	.166	.146	.224
Observations	10,645,769	9,185,412	9,511,130	7,080,688
Persons	3,351,593	3,301,936	3,097,049	2,347,598
Establishments	8,151	18,518	19,989	17,684

DATA: LIAB-AKM. NOTE: Daily wage (w) converted to log 1995 euros and top-coded values imputed. Time-varying effects $x'\beta$ not provided by IAB, and are estimated through a regression of $w - \pi - \phi$ on x. Years weighted equally in calculations.

The variance of full-time wages rose sharply over 1993-2017, from .169 to .232. Nearly all of the increase was attributable to the variances of π and ϕ and their covariance, with contributions of 52%, 14%, and 28%. The variance of the AKM residual rose only in proportion to $Var(w)$, while moments involving the time-varying effects $x'\beta$ had substantial but offsetting effects.[18] As a percentage of total wage variance, wage sorting (*i.e.* $Cov(\pi, \phi)$) rose from 9.5% to 14.6% over the sample period, and by the 2010-2017 panel was of comparable importance to firm-pay dispersion $Var(\phi)$.

Figure 3.1 plots the rise in wage sorting over the sample period, as both a covariance and a correlation, and with the corresponding (males-only) trend from Card et al. shown for reference. I am able to replicate the results from that paper, with one difference: I find $Corr(\pi, \phi)$ to be smaller and more volatile over time. This likely reflects the IAB's inclusion of women when estimating (4). West German women are more likely to work in small service sector establishments that, due to their size, have fewer of the job-movers needed for identification.[19] Including both sexes allows π and ϕ to be identified for a greater share of these employers. Relative to Card et al., the identified sample will skew more heavily towards services, an important difference given the sectoral differences characterized in Section 4; and it will skew towards small establishments, resulting in a larger incidence of (time-varying) limited-mobility bias, which I discuss below.

The rise in German wage sorting is robust in a number of respects. Results in Appendix B.2 show $Cov(\pi, \phi)$ to be positive and increasing across demographic groups and establishment characteristics. Wage sorting is also present in East Germany, though at initially greater levels and with a shallower increase over time. While the volatility evident in Figure 3.1 raises concerns regarding measurement and interpretation of the trend, I show in Appendix B.3 that among larger and better-identified establishments, the rise in wage sorting is approximately linear. The between-group results in the next section, which I establish below to be robust to statistical bias, further support the contention that the rise in $Cov(\pi, \phi)$ reflects real changes to the German economy.

Interpretation of Figure 3.1 is nevertheless challenging, for reasons that directly concern this paper.[20] First, π and ϕ are statistical objects, and even in a simple environment such as

[18] As the IAB provides only the estimates for π and ϕ, I obtain the time-varying effects through regressions of $w - \pi - \phi$ on x. As this step introduces additional error, the moments $Var(x'\beta)$, $Cov(\pi, x'\beta)$, and $Cov(\phi, x'\beta)$ should be interpreted cautiously.

[19] Wage effects can only be identified for a "connected set" of establishments that are linked through job-movers: individuals who transition between employers, but remain within the sample. As sample size falls, it becomes more likely that job transitions result in exit from the sample, and therefore the connected set shrinks.

[20] There are, in addition, challenges on which this paper is silent. Wage earnings form an incomplete picture of total compensation, and are subject to measurement issues and top-coding as discussed above. The standard approach of estimating (4) for full-time employees, while ensuring that only similar job spells are compared across time, necessarily leads to the omission of a sizable portion of the workforce. These are limitations common to AKM-based studies, which capture at best an incomplete picture of the sources of wage variation.

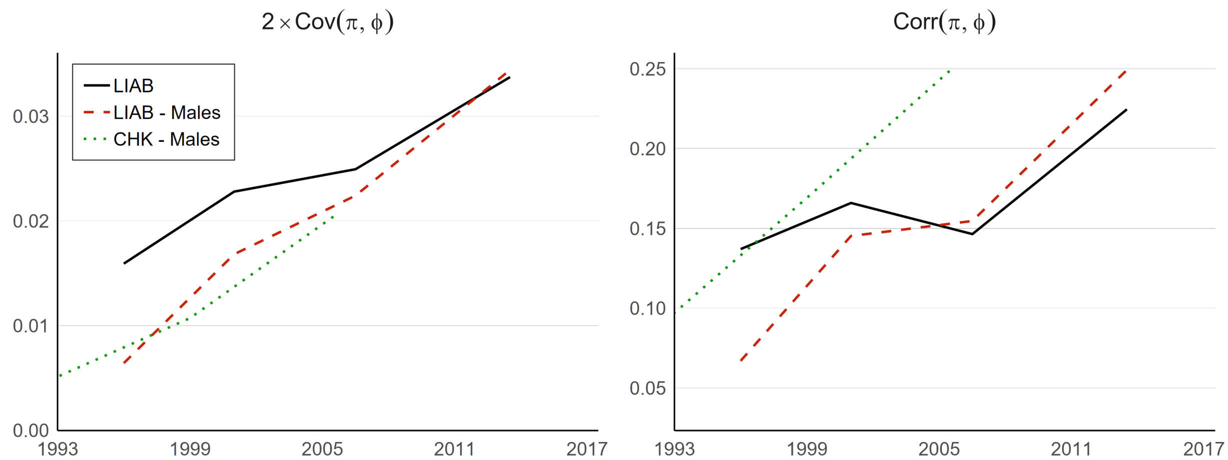

Figure 3.1: West German Wage Sorting, 1993-2017

DATA: LIAB-AKM and Card et al. (2013) Table 3. NOTE: Covariance and correlation of π and ϕ over AKM panel midpoint. Years weighted equally in calculations.

that considered in Section 2, there is no direct mapping between wage effects and theoretical sources of wage dispersion. This is particularly the case when match effects are present - a subject that I revisit in Section 4. Second, measures of $Cov(\pi, \phi)$ and $Corr(\pi, \phi)$ are subject to statistical bias, potentially confounding the analyses conducted in this paper.

Use of the IAB-provided wage effects directly addresses these challenges. Because (4) is implemented on a sample comprised of the universe of full-time job spells, the resulting estimates of π and ϕ are "best-case" in two respects. First, the loss of sample due to non-identification is minimized. This is important because identification is intrinsically related to establishment characteristics such as size, and is therefore non-random. A larger sample size also allows me to study group-level outcomes at a more dis-aggregated level. Second, estimation error is minimized, which in turn reduces the degree to which measures of wage sorting are statistically biased. Given the importance of this issue, I turn to it next.

3.3 Limited-Mobility Bias

A key challenge for this paper is that measures of wage sorting are subject to "limited-mobility" bias,[21] which arises due to the additive form of regression equation (4). Estimation errors associated with π and ϕ will be negatively correlated, biasing $Cov(\pi, \phi)$ and $Corr(\pi, \phi)$ downwards. As identification rests upon individuals who move between employers, the problem is worse for establishments with few entering or exiting workers (*e.g.* small employers), or when these transitions are not observed (*e.g.* due to limited sample coverage).

[21] See *e.g.* Andrews et al. (2008) and Bonhomme et al. (2023).

While limited-mobility bias is not observable, the identification statistics in Table 3.2 indicate that it is both materially present, and time-varying in the LIAB. Although the AKM wage effects are identified for the vast majority (97%-98%) of observations, sample loss is noticeably greater among small establishments, and for the 1998-2004 and 2003-2010 panels.[22] At establishments with less than 10 full-time employees, wage effects are unidentified for one-tenth of observations, rising to 13% in 2003-2010. Consistent with a greater incidence of bias, $Corr(\pi, \phi)$ is both lower for these establishments, and more variable over time. Among large establishments, the wage effects correlation rises monotonically over time and the deviation from trend in Figure 3.1 is entirely absent, suggesting that this is largely the product of time-varying bias.[23]

Table 3.2: AKM Identification Statistics

	1993-99	1998-04	2003-10	2010-17
		A. Unidentified Sample (%)		
All Establishments	2.2	2.8	3.0	2.3
... 1-9 Employees	10.4	12.4	13.3	11.8
... 10-24 Employees	0.5	0.6	1.3	0.5
... 25-99 Employees	0.4	0.5	0.4	0.3
... 100-499 Employees	0.4	0.4	0.5	0.4
		B. Wage Sorting ($Corr(\pi, \phi)$)		
1-9 Employees	-.091	-.100	-.212	-.112
10-24 Employees	.103	.079	.034	.206
25-99 Employees	.200	.228	.238	.299
100-499 Employees	.121	.227	.243	.301

Data: LIAB-AKM. Note: Panel A is the survey-weighted percent of observations with missing (unidentified) IAB-estimated AKM wage effects, and Panel B is the wage effects correlation. Establishment size defined over full-time workers.

The key advantage of the IAB-provided wage effects is that they are estimated over the universe of employment spells, and therefore sample loss and limited-mobility bias are minimized,[24] however the trade-off is that I am unable to implement the bias-correction estimator of Kline et al. (2020). In principle this estimator allows one to obtain unbiased measures of $Cov(\pi, \phi)$, but it would require access to the estimation data and the full set of regression effects, which are not made available by the IAB. It is important to note, in

[22] In Appendix B.3 I show that establishment entry and exit rates of full-time workers are lower during the 1998-2004 and 2003-2010 panels, and in particular during the latter of the two panels. This suggests that a decline in transition rates, and a resulting fall in the number of job-movers, is the root cause of the poor AKM identification during these panels.

[23] In Appendix B.3 I show similar results based on the number of job-movers per establishment, with the deviation from trend either small or absent among establishments with 20 or more job-movers.

[24] In a larger sample it will be more likely that one observes both employers associated with a job transition; hence limited-mobility bias is reduced by increasing the coverage of the sample.

addition, that the Kline et al. estimator introduces an *additional* bias when wage sorting is related to establishment size - which I find.[25] Implementation requires that establishments with a single job-mover be dropped. As these are overwhelmingly small establishments, the size distribution is distorted as a result.

While I cannot directly quantify the magnitude of limited-mobility bias in the LIAB, I am able to establish that the main results in this paper - which concern *group*-level outcomes - are robust to bias:

Proposition 3 (Unbiased Between-Group Covariance)**.** *Let G define a partition of a sample with N observations, where N_g is employment in group $g \in G$ and $N_{f,g}$ the corresponding measure for firm f in group g. Let π_g and ϕ_g denote population-mean values of π and ϕ, with $\hat{\pi}_g$ and $\hat{\phi}_g$ the corresponding estimates. If $\lim_{N \to 0} \frac{N_{f,g}}{N^g} = 0$ for all f and g, then $Cov(\hat{\pi}_g, \hat{\phi}_g)$ is unbiased at the limit.*

Proof. See Appendix A.3.

Given a large enough sample, the law of large numbers will hold at the group level, implying that group-mean values of π and ϕ are consistently estimated, and their covariance unbiased.[26] Note that while the condition $\lim_{N \to 0} \frac{N_{f,g}}{N^g} = 0$ will fail to hold for groups dominated by a small number of large establishments, these are cases where π and ϕ will themselves be consistently estimated, and limited-mobility bias absent.

Therefore the main implication of limited-mobility bias for this paper is that it may confound the within-group results in Section 4, and prevent me from ruling out a positive assortative matching mechanism. While the presence or absence of such a mechanism is of intrinsic interest, and I therefore provide a number of additional results in that section, the scope of my research question is limited to the *observed* wage sorting in Figure 3.1. As I find this to occur entirely at a group level, Proposition 3 provides a sufficient basis for obtaining bias-free results on this paper's object of study.

[25] See Appendix Table B.11. Roughly half of $Cov(\pi, \phi)$ occurs between employer size groups, and half within them. This is closely related to the between-industry results shown in the next section.

[26] This result is distinct from the observation in Bonhomme et al. (2019) that clustering firms prior to estimation reduces bias. Group-level means of AKM wage effects are not, in general, equivalent to group-level fixed effects from a wage regression. They will only be so when job mobility is random - a much stronger assumption than the "endogenous mobility" restriction imposed by the AKM model (see Card et al., 2013). Hence the argument in concerns a different, though related, set of outcomes.

4 Evaluating Sorting Mechanisms

In this section I evaluate the wage sorting mechanisms characterized in Section 2. Finding support for *technical sorting*, I close with a set of tests that further rule out *positive assortative matching* and *sorting on unobservables* as the relevant mechanisms for the wage sorting observed in Figure 3.1.

4.1 Group Covariance Decompositions

I begin with a set of decompositions that evaluate the consistency of German wage sorting with the between- and within-group predictions in Proposition 2. To this end I make use of the law of total covariance: given a partition of the sample $G = \{1, ..., N^g\}$, any covariance can be decomposed into the average within-group variance, plus the covariance of the group means:[27]

$$Cov(\pi, \phi) = \underbrace{\sum_{g \in G} \omega_g \Big(\pi - \mathbb{E}_g[\pi]\Big)\Big(\phi - \mathbb{E}_g[\phi]\Big)}_{\text{Within-Group Covariance}} + \underbrace{\sum_{g \in G} \omega_g \Big(\mathbb{E}_g[\pi] - \mathbb{E}[\pi]\Big)\Big(\mathbb{E}_g[\phi] - \mathbb{E}[\phi]\Big)}_{\text{Between-Group Covariance}}, \quad (6)$$

where $\mathbb{E}_g$ and $\mathbb{E}$ are the group-conditional and the unconditional expectations, and ω_g is the employment share of group g. Proposition 3 establishes that limited-mobility bias will be absent from the between-group covariance so long as N^g is not too large, with (6) implying that this bias will necessarily manifest in the within-group covariance.

A necessary step is to choose empirical analogues of output markets (z) and labor types (s). With respect output market segmentation, industry and location are natural choices. Although the version of the LIAB used in this paper only location only at the state (*Bundesland*) level, in Appendix B.4. I show that results are not substantively affected by augmenting state with additional variables that proxy for urbanicity.[28] To capture differentiation in labor markets, I consider both occupation and educational attainment, with the latter classified as in Card et al. (2013). Note, however, that two-thirds of individuals have the same reported attainment ('completed apprenticeship').

The between-group components, given in Table 4.1, show that observed wage sorting is predominantly between-industry and between-occupation, and *entirely* between industry-

[27]Variances can be similarly decomposed, allowing one to straightforwardly calculate a between-group correlation.

[28]In particular, I find that characteristics associated with urbanization (*e.g. proximity to university*) are more successful than Bundesland at explaining $Cov(\pi, \phi)$, consistent with the findings of Dauth et al. (2022) who find a substantial urban component to wage sorting; however I also find that the marginal explanatory power of these characteristics is negligible when controlling for industry and occupation, as urban-rural pay gaps are largely explanatory in terms of sectoral composition.

Table 4.1: Between-Group Wage Sorting

Grouping Variable	A. Covariance (% Total)				B. Correlation			
	93-99	98-04	03-10	10-17	93-99	98-04	03-10	10-17
Between-Output Markets (z)								
Industry (*12 groups*)	84.5	82.3	100.2	82.9	.747	.755	.773	.818
...Detailed (*46 groups*)	88.4	85.3	101.8	84.3	.674	.703	.723	.782
State (*11 Bundesland*)	6.7	3.7	5.1	2.7	.701	.500	.678	.495
Industry × State	89.3	85.2	103.2	82.7	.691	.725	.747	.744
Between-Labor Markets (s)								
Occupation (*15 groups*)	94.5	84.1	98.1	73.4	.590	.633	.646	.705
...Detailed (*75 groups*)	88.5	81.2	95.8	74.6	.462	.514	.534	.628
Education (*5 groups*)	24.4	27.5	35.3	35.6	.819	.940	.881	.830
Occupation × Education	96.2	87.6	102.2	79.8	.555	.612	.623	.672
Between-Both (z, s)								
Industry × Occupation	115.5	106.6	128.5	103.5	.476	.517	.549	.631
...Detailed	118.7	108.6	130.0	107.7	.402	.446	.479	.568

DATA: LIAB-AKM. NOTE: Covariance is the between-group component of $Cov(\pi, \phi)$ as a percent of total sample covariance. Correlation is the between-group covariance divided by the product of the between-group standard deviations. Years weighted equally in calculations.

occupation pairs.[29] Three-fourths of $Cov(\pi, \phi)$ occurs across aggregate occupational groups, and four-fifths across industries. When interacting occupation and industry, the between-group covariance is greater than the total, indicating a negative within-group component. On the other hand state is found to be uninformative, and education to be largely redundant with occupation. Though over-fitting is a concern in an exercise of this nature, it would not generate high between-group correlations (Panel B), which we also observe;[30] and results are largely unaffected by using highly-aggregated codes.

The results in Table 4.1 are not definitive, however, for reasons that I address in the remainder of the section. First, if the assumption of free entry imposed in Section 2 fails to hold, then *any* of the proposed mechanisms could conceivably generate between-group covariance. Second, while the positive within-industry and within-occupation covariances may indicate sorting on unobservables, they may also reflect technical sorting not captured by industry and occupation codes. The greater explanatory value of occupation × education provides some evidence in favor of the second explanation. And third, the negative covariance

[29]For ease of interpretation, the between-group covariance is reported as a percentage of the total covariance. While this measure is not robust to limited-mobility bias as it involves division by $Cov(\pi, \phi)$, it is appropriate for the exercise, the purpose of which is to attribute *observed* wage sorting to one (or some combination) of the candidate mechanisms.

[30]These reflect the fact, shown in Appendix B.4, that the preponderance of $Var(\pi)$ and $Var(\phi)$ is within industry-occupation cells.

within industry × occupation cells might simply reflect the influence of limited-mobility bias. Consistent with this explanation, within-group covariance is most negative during the 2003-2010 panel, where identification of the AKM effects is weakest. Therefore I next consider a set direct tests for the different mechanisms.

4.2 Testing For Technical Sorting

To test directly for technical sorting, I make use of the fact that the conceptual framework in Section 2 attributes this mechanism to price differentials arising from (fixed) entry costs. The prominent role of industry and occupation in Table 4.1 helps to constrain the set of relevant costs. One possibility is upfront capital investments, which are likely to vary across industries. Sectoral bargaining agreements will exert a similar affect, as while these would presumably affect the rent-sharing parameter ω rather than K, the two terms enter identically into the wage function. Across occupations, the most important cost differential is likely to be educational investments. Finally, wage differentials arising from non-monetary barriers to entry may also be consistent with technical sorting, if these allow incumbents to earn a greater match surplus than would be realized by a new entrant.

To test whether entry costs are predictive of the group-level variance and covariance of the AKM effects, I perform a first-stage regression that decomposes these into predicted and residual terms:

$$\pi_g = \delta x_g + \chi$$
$$\phi_g = \gamma z_g + \nu \ ,$$

where g subscripts denote group-mean values, and x and z are person and firm characteristics, respectively. Regression coefficients are reported in Panel A of Table 4.2, with groups defined as industry (12) × occupation (15) cells. I then perform a second-stage decomposition of the between-group covariance into predicted and residual components:

$$\begin{aligned} Cov(\phi_g, \pi_g) &= Cov(\delta x, \pi_g) + Cov(\chi, \pi_g) \\ &= Cov(\phi_g, \gamma z) + Cov(\phi_g, \nu) \ . \end{aligned}$$

These are reported in Panel B of Table 4.2. Results are based on the 2003-2010 LIAB-AKM panel, which has the broadest coverage of the relevant survey questions.

Regression results (Panel A) indicate that industry-occupation variation in the AKM wage effects is strongly predicted by observable measures, as evidenced by high R^2 values. Three-quarters of the variation in π is associated with education, and four-fifths of the

Table 4.2: Wage Sorting and Observable Heterogeneity, 2003-2010

	DV: Mean Worker Effect π_g				DV: Mean Firm Effect ϕ_g				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					A. Regression Results				
Years: Education	.154***		.162***	.103***					
	(.009)		(.007)	(.018)					
Years: Working		.018	.013**	.004					
		(.017)	(.007)	(.004)					
Years: At Firm		.005	.021***	.015***					
		(.013)	(.005)	(.003)					
$\log \frac{\text{Total Investment}}{\text{Employees}}$					.066***			.053***	.005
					(.009)			(.012)	(.011)
$\log \frac{\text{IT Investment}}{\text{Employees}}$					.043***			.084***	.044***
					(.012)			(.016)	(.017)
Bargaining Agm. (%)						.335***		.275***	.046
						(.102)		(.081)	(.047)
Competitive Mkt. (%)						.244***		.192***	.038
						(.064)		(.072)	(.084)
Log Employment							.090***	−.040**	.017
							(.007)	(.017)	(.011)
Earned Profits (%)							.565***	−.213	.084
							(.155)	(.182)	(.154)
R^2	.760	.040	.860	.960	.790	.180	.600	.830	.940
					B. Predicted Covariance				
Predicted (%)	.430	.250	.920	.950	.840	-.010	.670	.850	.900
Residual (%)	.570	.750	.080	.050	.160	1.01	.330	.150	.100

DATA: LIAB-AKM. NOTE: Weighted results across 179 industry-occupation cells (1 omitted due to insufficient observations). Bold text indicates statistical significance at the 5% level; robust standard errors shown. See Appendix Table B.2 for variable definitions.

variation in ϕ with capital-intensity (measured as investment *per capita*).[31] Conditional on these, the remaining variables provide only marginal explanatory power. Within-group regressions (specifications 4 and 9) yield qualitatively similar coefficients, and are indicative of variation in K and in C that is not captured by industry and occupation codes, respectively.

The second-stage decomposition (Panel B) shows that barriers to entry are similarly predictive of the between-group covariance: $Cov(\delta x, \pi_g)$ and $Cov(\phi_g, \gamma z)$ account for 92% and 85% of the total, respectively. They are, in fact, more predictive than industry or occupation codes, which by themselves predict less than 80% of the between-group covariance. Capital

[31] I use investment as capital stocks are not observed, and include information technology as a regressor as it is more likely to correlate with software and other forms of unobserved intangible capital (see Brynjolfsson et al., 2021). Collectively bargained wage agreements are included as they are widespread In Germany, notably among manufacturing establishments. Years of education are imputed as in Card et al., while labor force experience is calculated the number of years since first paying social security taxes. See Appendix B.1 for additional details on variables and formatting.

investment is sufficient to drive the result for $Cov(\phi_g, \gamma z)$, however education and experience both contribute to the high value of $Cov(\delta x, \pi_g)$. While a role for experience and/or job tenure is consistent with technical sorting as argued above, it may also reflect selection of sorted matches, consistent with an assortative matching mechanism. I revisit (and rule out) this possibility below, and I provide additional results on experience's role in Section 6.

The results of this test strongly support a technical sorting mechanism, in two respects. First, measures of explicit and implicit entry costs are sufficient to predict all of the between-industry and between-occupation covariances in Table 4.1. The two most explanatory variables are education and capital costs, which are directly analogous to C and K in the conceptual model. Second, much of the within-industry and within-occupation covariances are *also* predicted by the variables in Table 4.2. Although the conceptual framework would attribute these variance components to sorting on unobservables (ability or productivity), it is likely that to some extent they reflect technical sorting not captured by industry and occupation codes.

4.3 Testing For Assortative Matching

I supplement the previous result with a direct test for assortative matching, which also serves to test for any sorting on unobservable characteristics that is subject to search frictions. A central prediction of such mechanisms is *dynamic selection*: when the other agent's type - and therefore match effect $\hat{y}$ - are unknown *ex ante*, any sorting must arise through the destruction of non-sorted matches.[32] If dynamic selection is the source of wage sorting, then we would expect it to be weakest among new matches and newly-entered workers and establishments. Note that while this logic assumes search to be frictional, some such assumption is needed in order for assortative matching to generate wage sorting, as otherwise free entry will tend to price out match effects as discussed in Section 2.[33]

Figure 4.1 shows that in fact, wage sorting is *strongest* among new matches, new workers, and new establishments. This result is the opposite of what we would expect if the sorting patterns driving $Cov(\pi, \phi)$ were impeded by search frictions. Importantly, limited-mobility bias does not explain the patterns seen in the Figure: all variables are positively associated

[32] In the stylized framework of Shimer and Smith (2000), match selection occurs immediately through agents' acceptance sets. In practice we would expect learning to occur over time, and match selection to occur dynamically through on-the-job search, terminations, and/or employer exit. For example, in the environment of Lentz (2010) and Bagger and Lentz (2019) sorting patterns strengthen as workers move up the job ladder, and are weakest among those transitioning from unemployment.

[33] This argument applies most forcefully to assortativity occurring between groups (industry or occupation), which is the target of this exercise given the results in Table 4.1. Across industries, higher match effects will induce entry and lower output prices, whereas an occupation with high match effects will induce a change in relative shares that - under standard assumptions regarding the production technology - lowers marginal product.

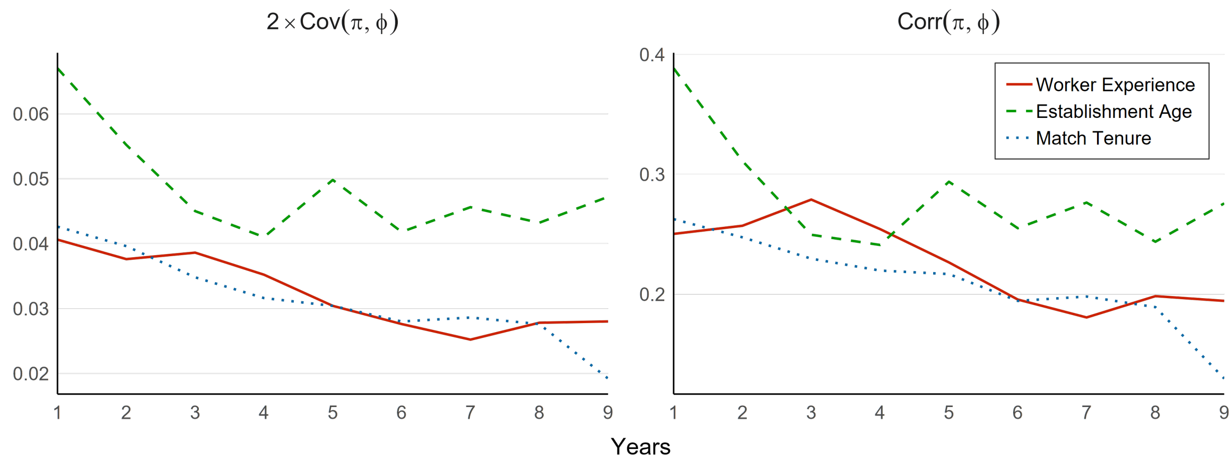

Figure 4.1: Wage Sorting and Match Selection, 2010-2017

DATA: LIAB-AKM. NOTE: Wage sorting measures conditional on years of experience, age, and tenure. Worker experience dated from first payment into social security. Match tenure is years employed at current establishment.

with the number of job-movers per establishment,[34] and therefore bias would tend to work against the negative relationships observed. Though inconsistent with assortative matching, a high value of $Cov(\pi, \phi)$ among new matches and entrants is easily explained in terms of sectoral composition. Service industries are over-represented among these groups, and as π and ϕ are more correlated within services than in the aggregate, a higher covariance results.

In the Appendix I present a pair of supplementary results. First I show that under various specifications, the relative wage gain from moving to a higher-paying firm is slightly *smaller* for high-earning workers. If wage sorting were due to match complementarities, as implied by assortative matching, then we would expect these moves to yield a relatively larger increase in match output and hence wages. Second, I perform the group decomposition in Table 4.1 for a "well-identified" sample of establishments with 20 or more job-movers, for which limited-mobility bias should be largely absent. While for this sample the wage effects are positively correlated within industry-occupation cells, the correlation is small ($\sim .06$). Industry and occupation continue to jointly account for 80-90% of the total covariance and 90-94% of the 1993-2017 trend, with the latter number roughly unchanged from Table 4.1.

5 Explaining Wage Sorting's Rise

In this section I decompose the rise in wage sorting over the 1993-2017 period into three sources of change: industry-occupation employment shares, group-mean person AKM effects,

[34]See Appendix B.4.

and group-mean firm AKM effects. Finding that the first two explain the bulk of the trend, I then provide evidence linking these changes to a general rise in the skill premium and domestic outsourcing to temporary employment agencies.

5.1 Trend Decomposition

As German wage sorting occurs entirely at the industry-occupation level, the 1993-2017 trend can only have arisen from changes to the distribution of employment, and/or the distributions of AKM wage effects across industry-occupation pairs. Starting with equation (6) and letting hat variables indicate demeaned values, we can decompose the change in $Cov(\pi_g, \phi_g)$ into the individual contributions of ω_g, π_g, and ϕ_g, plus a set of interaction terms, omitted here for brevity:

$$\begin{aligned} Cov(\pi'_g, \phi'_g) - Cov(\pi_g, \phi_g) &= \sum_{g\in G} \left[\hat{\omega}'_g \hat{\pi}'_g \hat{\phi}'_g - \hat{\omega}_g \hat{\pi}_g \hat{\phi}_g\right] \\ &= \underbrace{\sum_{g\in G}\left[\left(\hat{\omega}'_g - \hat{\omega}_g\right)\frac{\hat{\pi}'_g\hat{\phi}'_g + \hat{\pi}_g\hat{\phi}_g}{2}\right]}_{\text{Individual Contribution, } \omega_g} + \underbrace{\sum_{g\in G}\left[\left(\hat{\pi}'_g - \hat{\pi}_g\right)\frac{\hat{\omega}'_g\hat{\phi}'_g + \hat{\omega}_g\hat{\phi}_g}{2}\right]}_{\text{Individual Contribution, } \pi_g} \\ &\quad + \underbrace{\sum_{g\in G}\left[\left(\hat{\phi}'_g - \hat{\phi}_g\right)\frac{\hat{\omega}'_g\hat{\pi}'_g + \hat{\omega}_g\hat{\pi}_g}{2}\right]}_{\text{Individual Contribution, } \phi_g} + \mathit{Interactions}\ . \end{aligned} \tag{7}$$

A shift in ω_g could reflect technical or structural change (*i.e.* sectoral or occupational demand shifts), or increasing segregation of high-earning occupations into high-paying sectors. Changes to the distribution of π_g could be driven by skill premia, or by changes to the manner in which workers sort across jobs. Values of ϕ_g may be affected by trends in collective bargaining or output market competition. While decomposition (7) cannot positively identify the mechanism(s) responsible for wage sorting's rise, it can serve to rule out those that are least likely to have played a role.

Column 4 of Table 5.1) shows that changes to industry employment shares and industry-mean person effects account for substantively all of the 1993-2017 trend. Interaction terms are negligible in the aggregate, with the individual components summing to 98% of the total change. Of this, nearly all is due to changes in π and ω. Although the variance of ϕ increases over the first half of the sample, thereby raising $Cov(\pi, \phi)$, this development reverses during the 2010's and therefore the effect over the full sample period is small.[35] On the other hand

[35] At a disaggregated level I observe two offsetting trends. Consistent with arguments made in Card et al. (2013), I find a rise in the within-industry dispersion of ϕ that appears related to declines in collective bargaining coverage

Table 5.1: Decomposition of Trend

	1993-99 - 1998-04	1998-04 - 2003-10	2003-10 - 2010-17	1993-99 - 2010-17
Total Between-Group	0.0030	0.0038	0.0014	0.0082
Individual Contribution (%)				
Employment Share ω	42.0	55.3	127.8	52.1
... Between-Industry	39.4	44.9	121.9	49.1
... Between-Occupation	7.2	11.5	10.8	7.6
Person Effect π	27.0	10.5	159.0	41.2
... Between-Industry	30.2	20.5	153.5	43.0
... Between-Occupation	13.1	4.3	57.1	16.1
Employer Effect ϕ	31.2	34.3	-188.8	4.5
... Between-Industry	22.6	19.7	-172.9	-5.2
... Between-Occupation	51.0	53.6	-59.4	34.9
Industry-Occupation Pairs	176	178	179	176

DATA: LIAB-AKM. NOTE: 'Total between-group' is the difference in $Cov(\pi_g, \phi_g)$ across panels. Individual contributions from (7) are estimated across 180 industry-occupation pairs and reported as a percent of the total change. Between-group results hold fixed the within-group distribution; see Appendix B.5 for details.

the relevant changes to π and ω occurred at the industry level, as their contributions are mostly unaffected by holding fixed their within-industry distributions (see 'between-industry' results). This would seem to rule out occupational segregation or similar explanations.

To provide intuition as to the underlying changes taking place, I plot the industry- and occupation-mean wage effects in Figure 5.1 for the first and last AKM panels. At the industry level, we observe a general widening of π-differentials in the 2010-2017 panel, while ϕ differentials are largely similar to their 1993-2017 values. We see as well a shift in employment from materials and crafts manufacturing to personal services, hospitality, and temp agencies.[36] While these industries are relatively similar in terms of worker earnings, employer pay is 20% higher in manufacturing. Hence the shift in employment serves to strengthen the overall correlation of π and ϕ. Across occupations, this change manifests as a downward movement (decrease in ϕ) among lower-skilled manual labor trades, as these become more concentrated in services; and an upward movement among other goods-producing trades, as manufacturing employment becomes increasingly concentrated in higher-paying industries. I revisit these changes below, when I quantify the industry- and occupation-level contributions.

among small employers. As employment in high-π occupations is increasing in firm size, the result is greater within-industry wage sorting. At the same time, ϕ has risen among low-π service establishments, which has tended to reduce overall wage sorting. These results are shown in Appendix B.6.

[36]Hospitality services and temp agencies are aggregated so as to preserve sample sizes, as these industries are relatively sparsely populated in the IAB survey for the 1993-1999 panel.

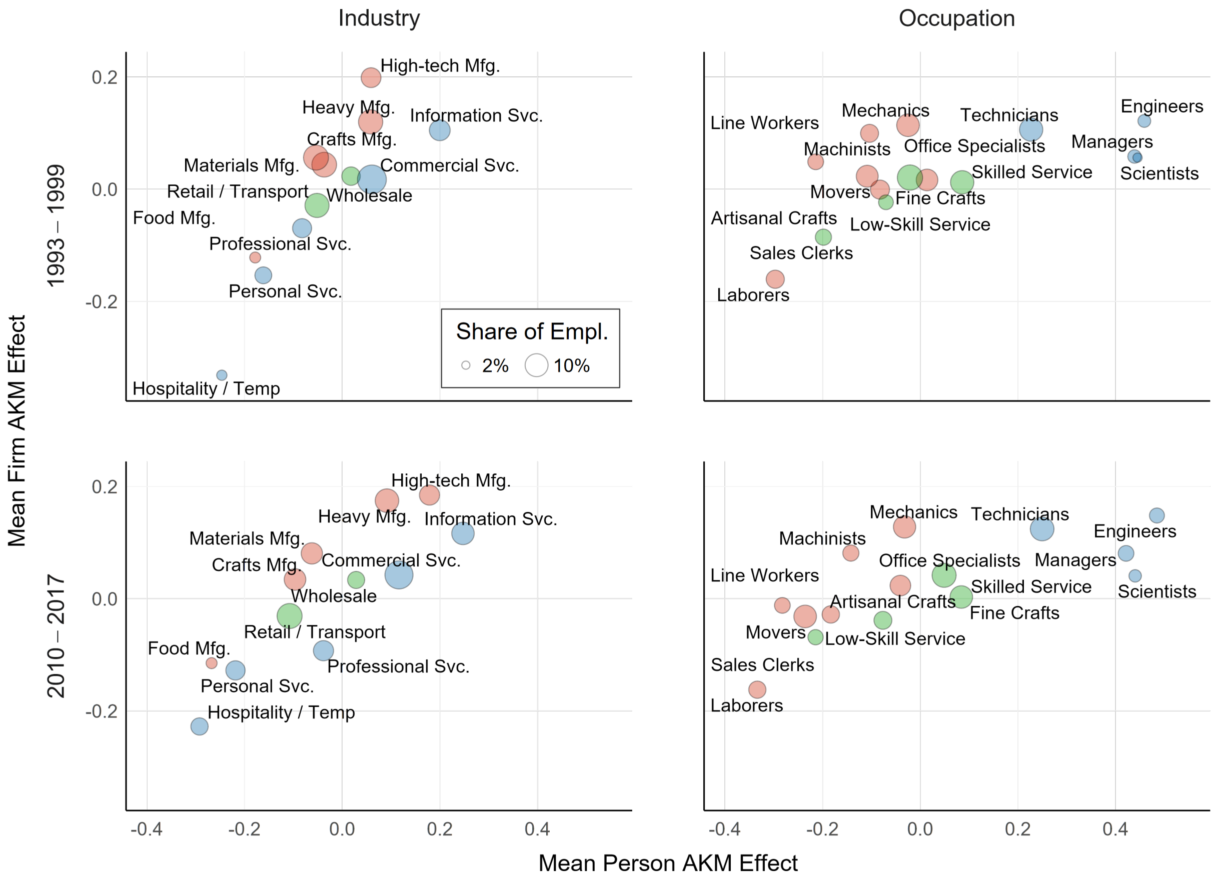

Figure 5.1: Group-Mean AKM Wage Effects

Data: LIAB-AKM. Note: Weighted average of AKM wage effects by aggregated WZ 2008 industry. Bubble size indicates share of total employment. Red indicates manufacturing industries (goods-production occupations), green commerce (sales & service), and blue services (professional).

5.2 Skill Premia Versus Sorting Patterns

A rise in mean worker effect π_g among high-paying industries may reflect either a general rise in the price of skill, or changes to the sorting of workers across jobs. The return to education has risen in Germany since the 1990's, though at a slower pace than in the United States citepDoepke2024, and if better-educated workers sort into higher-paying sectors as indicated in Table 4.2 then a rising education premium would naturally contribute to $Cov(\pi_g, \phi_g)$. On the other hand, it may be that high-paying sectors saw an increase in worker ability relative to low-paying sectors, manifesting either as an increase in worker education, or a rise in residual wages (*i.e.* controlling for education).

To distinguish these possibilities, I regress (industry × occupation) group-mean person effects π_g on years of education, and perform a modified Oaxaca-Blinder decomposition.

Writing the first-stage regression as

$$\pi_g = \lambda_0 + \lambda_{Ed} Ed_g + \chi_g \ , \tag{8}$$

where Ed_g is the mean of imputed years of education following Card et al., the term $\left(\hat{\pi}'_g - \hat{\pi}_g\right)$ in equation (7) may be decomposed as

$$\hat{\pi}'_g - \hat{\pi}_g = \left(\lambda'_{\mathrm{Ed}} - \lambda_{\mathrm{Ed}}\right) \frac{\hat{Ed}'_g + \hat{Ed}_g}{2} + \frac{\lambda'_{\mathrm{Ed}} + \lambda_{\mathrm{Ed}}}{2} \left(\hat{Ed}'_g - \hat{Ed}_g\right) + \left[\chi' - \chi\right] \ ,$$

with hats indicating demeaned variables as before. One may then separate the individual contribution of π_g into three components: changes to mean group education, changes to the education wage premium, and a residual term.

Table 5.2: Oaxaca-Blinder Decomposition - Contribution of π_g

	Total Contribution	Between-Industry	Between-Occupation
	A. Education Regression Coefficients		
$\lambda_{\mathrm{Ed}}^{1993-99}$	.150***	.162***	.155***
	(.010)	(.017)	(.010)
$\lambda_{\mathrm{Ed}}^{2010-17}$	.162***	.167***	.165***
	(.009)	(.007)	(.008)
	B. Individual Contribution of π_g (%)		
Total Contribution	41.2	43.0	16.1
... Change in λ_{Ed}	17.9	26.2	7.3
... Change in Ed	5.1	1.3	4.2
... Change in χ	18.2	15.5	4.6

DATA: LIAB-AKM. NOTE: Oaxaca-Blinder decomposition of individual contribution of π_g between the 1993-99 and 2010-17 panels, across 176 industry-occupation cells (4 omitted due to insufficient observations), 12 industry groups, and 15 occupation groups. Ed is mean within-group education and λ_{Ed} is the coefficient from regression (8). Bold text indicates statistical significance at the 5% level; robust standard errors shown. All results in percentage of total trend.

Results of the Oaxaca-Blinder decomposition, shown in Table 5.2, indicate that rising education premia are the main reason for the increase in wage sorting attributable to π_g. Group education differentials are roughly stable, and explain only a small part of the rise in $Var(\pi)$.[37] Of the between-industry contribution of π, three-fifths is predicted by the rise in λ_{Ed}, with most of the remainder corresponding to the residual term in (8). The unexplained portion may indicate greater sorting along unobserved ability (σ in the conceptual frame-

[37]Some notable up- or down-skilling is however present. Educational attainment rose most among skilled service occupations, technicians, and office specialists, whereas the smallest increases occurred among crafts and transportation occupations, and in science and engineering. This result is shown in Appendix B.7. I nevertheless find that industry and occupation education differentials change little over the sample period.

work), or simply between-industry variation in skill or human capital not captured by *Ed*. Overall, however, results point to an interaction between a rising skill premium and stable industry-education gaps.

5.3 Structural Change Versus Domestic Outsourcing

The contribution of ω_g, on the other hand, may reflect known patterns of structural change - for example, a rising service share as in the U.S. - or shifts in sectoral composition that are unique to Germany. To shed additional light, I begin by disaggregating the between-group results in Table 5.1 to the industry and occupation levels, which is easily done given the additive form of decomposition (7). Group-level contributions are plotted in Figure 5.2, together with the associated change in employment share. For this exercise I hold fixed the within-group distributions (for example, of occupational shares within a given industry), so as to isolate the impact of changes to group shares.

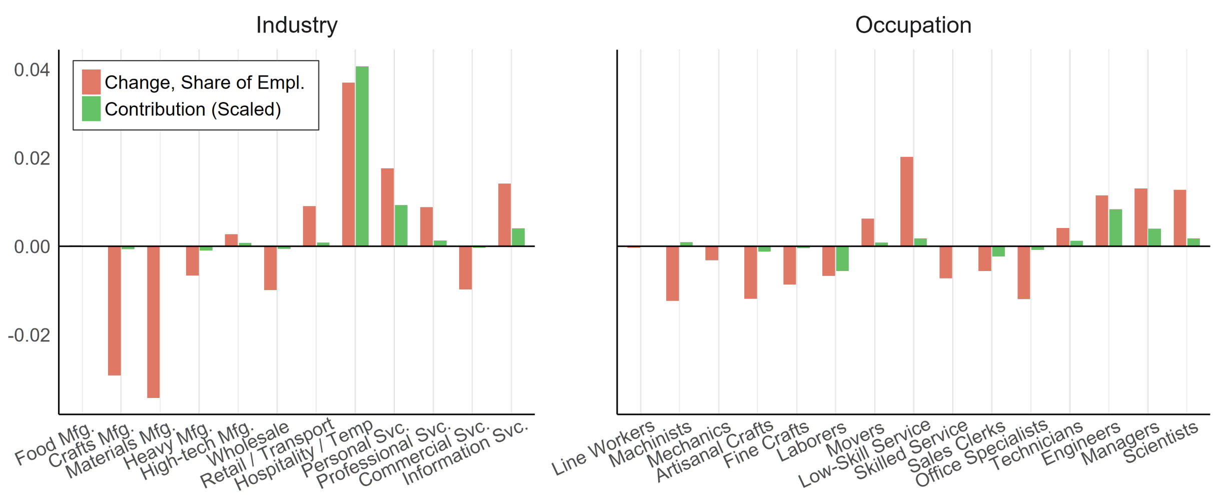

Figure 5.2: Contribution of ω_g - Industry and Occupation

DATA: LIAB-AKM. NOTE: Change in ω_g and individual contributions from decomposition (7) across 176 industry-occupation cells (4 omitted due to insufficient observations). Within-group shares held fixed. Changes to ω denoted as a percent of total employment. Individual contributions scaled.

Strikingly, four-fifths of the between-industry contribution is associated with growth of temporary employment agencies and hospitality establishments, with the former rising by two percentage points and the latter by a smaller amount.[38] Both increases are likely to be related to domestic outsourcing. Past studies associate increased temp labor with the

[38] Hospitality services and temp agencies are aggregated so as to preserve sample size in the 1990's, a period in which these industries are sparsely sampled in the IAB. Note, however, that this does not affect measured employment shares, which incorporate use of survey weights.

deregulation of the Hartz reforms (Spermann, 2011; Garz, 2013), and service sector growth will reflect any unit-level outsourcing as studied by Goldschmidt and Schmieder (2017). Therefore the results in the Figure suggest a quantitatively important link between wage sorting and domestic outsourcing, consistent with the intuition that a key incentive for such outsourcing is the reduction of wages.

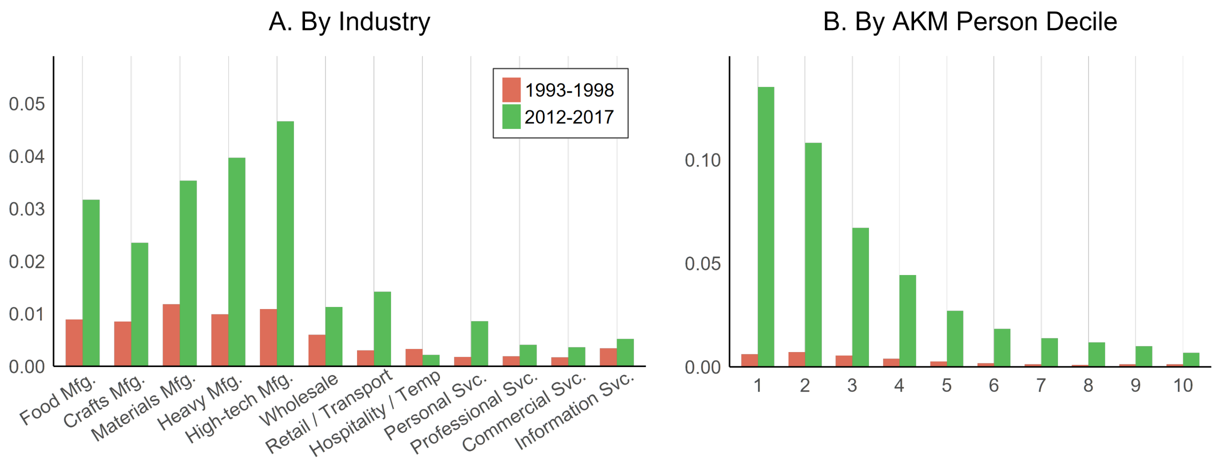

Figure 5.3: Temp Agency Employment

DATA: LIAB-AKM. NOTE: Temporary agency employment as percent of total workforce, by industry (Panel A) and by equal-weighted person AKM effect decile (Panel B). Panel B includes both temporary workers and staff of temporary agencies.

The mechanical relationship between temp employment and wage sorting is apparent from Figure 5.3, which plots the share of temp workers by industry and AKM person effect decile. Use of temp labor is concentrated in manufacturing, and is most prominent among the highest-paying industry groups (high-tech and heavy manufacturing). Temp workers on the other hand are overwhelmingly low-earning, with the vast majority coming from the lower half of the π-distribution. As temp agencies are low-paying, and temp workers low-earning, any growth in this sector would tend to increase $Cov(\pi, \phi)$. The effect will be particularly large, however, if temp workers are outsourced by manufacturing establishments, as this would translate into a roughly 30% decrease in pay.[39] Hence the *causal* effect of temp agency growth on wage sorting will depend on the extent to which it has displaced labor in German manufacturing - a question that lies outside the scope of this paper.

[39] In Appendix B.7 I show that workers who transition from the heavy or high-tech manufacturing sectors to temp agency employment see a roughly 30% pay decrease, while those who transition in the reverse direction see pay increases of similar, though slightly smaller magnitudes.

6 What Drives Technical Sorting?

If wage sorting is technical in nature as argued so far, then in the framework of Section 2 it must be that $Cov(K, C) > 0$ for some set of output market entry costs K and labor entry costs C. In this section I characterize the observable analogues of K and C, and I present a set of descriptive results on possible reasons for their covariance. These results are not intended to be conclusive, but to inform future research by addressing the question: what are the observable features of German wage sorting, and are they consistent with one or more general mechanisms that may account for wage sorting outside of Germany?

6.1 Characterizing K and C

To characterize the relationships between wage sorting and observable entry costs (explicit or implicit), I begin with standardized cross-sectional regressions of the form:

$$\hat{\phi}_g \sim \eta_\pi^{FE} \hat{\pi}_g + \eta_\pi^X \hat{X}_{\pi,g} + \psi_{\pi,g} \tag{9}$$

$$\hat{\pi}_g \sim \eta_\phi^{FE} \hat{\phi}_g + \eta_\phi^X \hat{X}_{\phi,g} + \omega_{\phi,g} \ , \tag{10}$$

where g indicates group means as before, X is a set of candidate entry costs, and hatted variables are Z-score normalized. In the absence of controls, η^{FE} will simply give the correlation between π_g and ϕ_g. If an cost variable X moderates this relationship, then we would expect (1) a large and positive value of η^X, consistent higher match surplus and therefore wages, and (2) a lower value of η^{FE} when X is included as a control. Coefficients from single regressions are given in the first two columns of Table 6.1.[40]

I then isolate the influence of each cost measure through a Gelbach decomposition, which consists of four steps. In the case of regression (9), these include: (1) obtaining η^{FE} when $X = \emptyset$, (2) obtaining η^X when X includes the full set of cost measures, (3) obtaining the coefficient from a simple regression of X on $\hat{\pi}_g$, and (4) taking the difference between η^{FE} from step 1 and the product of steps 2-3. This yields for each X a measure of that variable's contribution η^{FE}. Results from the joint regressions (step 2) are given in columns 3-4 of Table 6.1, and the final two columns give the results of the Gelbach decomposition.

Results in Table 6.1 support capital investment, and IT investment in particular, as the relevant measures for firm entry cost K.[41] The correlation of .78 with ϕ_g, given by η^X in the

[40] Variable definitions are given in Appendix B.1, and results for additional variables in Appendix B.8. One limitation of this analysis is that, because the majority of IAB survey variables are present only in a small subset of years, or are not fully comparable over time, the results in the table are cross-sectional. While a panel analysis would be of interest, the stability of the wage sorting patterns shown in Section 2 - and characterized further in this section - suggests that such an analysis is unlikely to yield different results.

[41] Note that IT is correlated with product development and formalized job reviews - proxies for knowledge and

Table 6.1: Standardized Regressions and Gelbach Decomposition, 2003-2010

	Single Regression		Joint Regression		Gelbach Decomp.	
	η^X	η^{FE}	η^X	η^{FE}	Contr.	% Expl.
A. Dependent Variable: Firm Effect ϕ_g						
No Control		.548***				
		(.086)				
Log $\frac{\text{Total Investment}}{\text{Employees}}$	.765***	.241***	.547***	.146***	.219***	40.0
	(.050)	(.051)	(.127)	(.048)	(.079)	
Log $\frac{\text{IT Investment}}{\text{Employees}}$	.782***	.042	.539***	.146***	.349***	63.7
	(.099)	(.069)	(.164)	(.048)	(.125)	
Product Development (%)	.670***	.262***	−.272*	.146***	−.116*	−21.2
	(.057)	(.064)	(.150)	(.048)	(.067)	
Written Perf. Reviews (%)	.492***	.299***	−.077	.146***	−.039	−7.1
	(.117)	(.099)	(.143)	(.048)	(.074)	
Competitive Market (%)	.329***	.584***	.219***	.146***	−.024	−4.4
	(.091)	(.074)	(.085)	(.048)	(.020)	
Positive Profits (%)	−.133	.582***	−.006	.146***	−.002	−0.3
	(.110)	(.091)	(.051)	(.048)	(.013)	
Bargaining Agrmnt. (%)	.299***	.527***	.203***	.146***	.014	2.6
	(.089)	(.088)	(.078)	(.048)	(.020)	
B. Dependent Variable: Worker Effect π_g						
No Control		.548***				
		(.080)				
Years: Education	.782***	.336***	.625***	.150**	.169***	30.9
	(.034)	(.031)	(.058)	(.061)	(.055)	
Knowledge: Math/Sci.	.532***	.404***	.161**	.150**	.100***	18.3
	(.071)	(.066)	(.068)	(.061)	(.033)	
Knowledge: Business	.576***	.563***	−.042	.150**	−.035	−6.4
	(.063)	(.041)	(.071)	(.061)	(.047)	
Knowledge: Technical	.051	.526***	.058	.150**	.043*	7.9
	(.078)	(.076)	(.075)	(.061)	(.025)	
Knowledge: PC Applic.	.701***	.238***	.211***	.150**	.001	0.2
	(.061)	(.052)	(.054)	(.061)	(.004)	
Years: Working	−.245**	.701***	.161***	.150**	.025	4.6
	(.105)	(.109)	(.048)	(.061)	(.032)	
Years: At Firm	−.670***	1.069***	−.045	.150**	.093***	17.0
	(.121)	(.110)	(.060)	(.061)	(.034)	

DATA: LIAB-AKM and 2006 BIBB Survey. NOTE: Results from weighted regressions across 179 industry-occupation cells (1 omitted due to insufficient observations). Variables Z-score normalized. η^{FE} and η^X are coefficients from (9) in Panel A, and (10) in Panel B. Joint regressions include all cost measures. Gelbach contribution is X's effect on η^{FE}, with final column giving each contribution as a percentage of $Corr(\pi_g, \phi_g)$. Bold text indicates statistical significance at the 5% level; robust standard errors shown.

organizational capital - which accounts for the negative and insignificant coefficients on these measures in the joint regression. Hence IT investment can be interpreted as, at least in part, a proxy for intangible capital. The primary difference between IT and total investment is that the latter is associated with manufacturing and hence the manufacturing pay premium - which, as discussed earlier, tends to dampen wage sorting.

single regression, indicates close to perfect correlation with group-mean establishment pay. Controlling for IT reduces the relationship between ϕ_g and π_g to statistical insignificance, and the Gelbach decomposition suggests that wage sorting would be stronger if variation in non-capital costs were absent. Measures of intangible and organizational capital (product development and formal performance reviews) are explanatory individually, but not in joint regressions, reflecting their strong correlation with IT investment. Competition in output markets (presumably indicating a lack of monopsony power) and bargaining agreements are consistently associated with ϕ_g, but orthogonal to $Cov(\pi_g, \phi_g)$.

With respect to labor entry cost C, results suggest a role for both human capital and worker experience. In both individual and joint regressions, imputed years of education and STEM-related domain knowledge are both associated with both π_g and $Cov(\pi_g, \phi_g)$. Notably, PC-related knowledge is predictive in single but not joint regressions, its explanatory power evidently due to a correlation with the two human capital measures just mentioned. Worker experience and job tenure, which are positively (and to some extent mechanically) correlated with each other, are negatively associated with worker earnings and hence with wage sorting. This however reflects a lower degree of education, with columns 3-6 indicating a positive contribution to wage sorting. I revisit this relationship below.

6.2 IT and Education

Table 6.1 and the results from Section 4 indicate that wage sorting is closely associated with an observable relationship, occurring at the sectoral level, between employer fixed capital (especially IT) and human capital. While these results do not establish causality, they are consistent with the logic of the conceptual framework. If capital investment is an upfront cost, then higher values of K and C will reduce entry, thereby raising match surplus and driving larger estimated person and employer wage effects.

But does the association between IT and human capital represent a general relationship? While the goal of this paper is to identify the drivers of $Cov(\pi, \phi)$, rather than the mechanism(s) responsible for a positive value of $Cov(K, C)$, this paper's results will be more meaningful if they shed light on wage sorting outside of Germany. In the case of information technology and education, there are several general mechanisms consistent with the existing literature. Therefore I describe these mechanisms briefly, and provide some preliminary evidence on them.

I differentiate between the following three mechanisms, which can be distinguished by their implications for the types of labor used in production:

1. **Technological skill-bias**. IT-intensive sectors are skill-intensive due to elimination

of routine jobs: $\frac{\text{ICT Investment}}{\text{Employees}}$ is associated with a larger share of non-routine labor.

2. **Organizational complexity**. IT-intensity and skill-intensity are each the result of organizational complexity: $\frac{\text{ICT Investment}}{\text{Employees}}$ is associated with a larger share of managerial and bureaucratic labor.

3. **Knowledge production**. Knowledge-intensive sectors are more IT-intensive and more skill-intensive: $\frac{\text{ICT Investment}}{\text{Employees}}$ is associated with a larger share of labor employed in knowledge production.

The first follows directly from the literature on routine-bias: information technology is widely supposed to reduce the demand for routine (and typically less-skilled) occupations. The second is consistent with theory and evidence linking capital-intensity to firm size (*e.g.* "superstar" firms as in Autor et al. 2020), and notions of knowledge hierarchies as in Garicano (2000), in which firms expand by adding layers of management. The third would arise naturally if IT and intangible capital are complements in the production of knowledge-intensive outputs (see Brynjolfsson et al., 2021), to the extent that intangible capital is the product of R&D and other skill-intensive activities performed within the firm.[42]

The predictions of these mechanisms are evaluated informally in Figure 6.1, with evidence supporting the third: IT investment is associated with a greater share of knowledge-intensive occupations and research-focused tasks. We would expect a skill-bias mechanism to less employment in routine production and sales occupations (Autor et al., 2003; Acemoglu and Autor, 2011), however we see a broadly flat relationship with respect to IT investment. Likewise, IT-intensive sectors exhibit similar shares of managers and office professionals, and a similar frequency of organizational tasks like "advising and informing colleagues". On the other hand, STEM jobs and tasks related to knowledge production are more common in these industries, offset by a lower share of manual labor occupations (unshown).[43] A set of Gelbach decompositions in Appendix B.8 quantitatively supports these results.[44]

[42]Note that wage sorting is not a necessary outcome of any of the three mechanisms. The effects of skill-bias will depend on within- and between-sector elasticities of substitution. Models of knowledge hierarchies are often ambiguous with respect to their relationship with the occupational structure of labor markets. Knowledge-intensity is likely to vary with other sectoral differences that may affect capital intensity and hence employer pay. Any mechanism would, in addition, need to be sufficiently strong as to substantively explain observed variation in K and C.

[43]While these results potentially consistent with outsourcing - and thus with findings in the previous section - I show in Appendix B.8 that establishment outsourcing is only weakly related to η^X and hence wage sorting. Notably, wage sorting is strongly present within the services sector, where outsourcing is mostly absent.

[44]I also show several auxiliary results inconsistent with the skill-bias or organizational hypotheses. In particular, I find that sectoral skilled shares and skill premia have evolved similarly for high- and low-IT industries, whereas we would expect a skill-bias mechanism to have strengthened over the 1993-2017, resulting in a response of either wages or shares. On the other hand skill premia are generally lower in high-IT sectors, whereas knowledge hierarchies predict a greater return-to-skill in complex organizations (Garicano and Rossi-Hansberg, 2015).

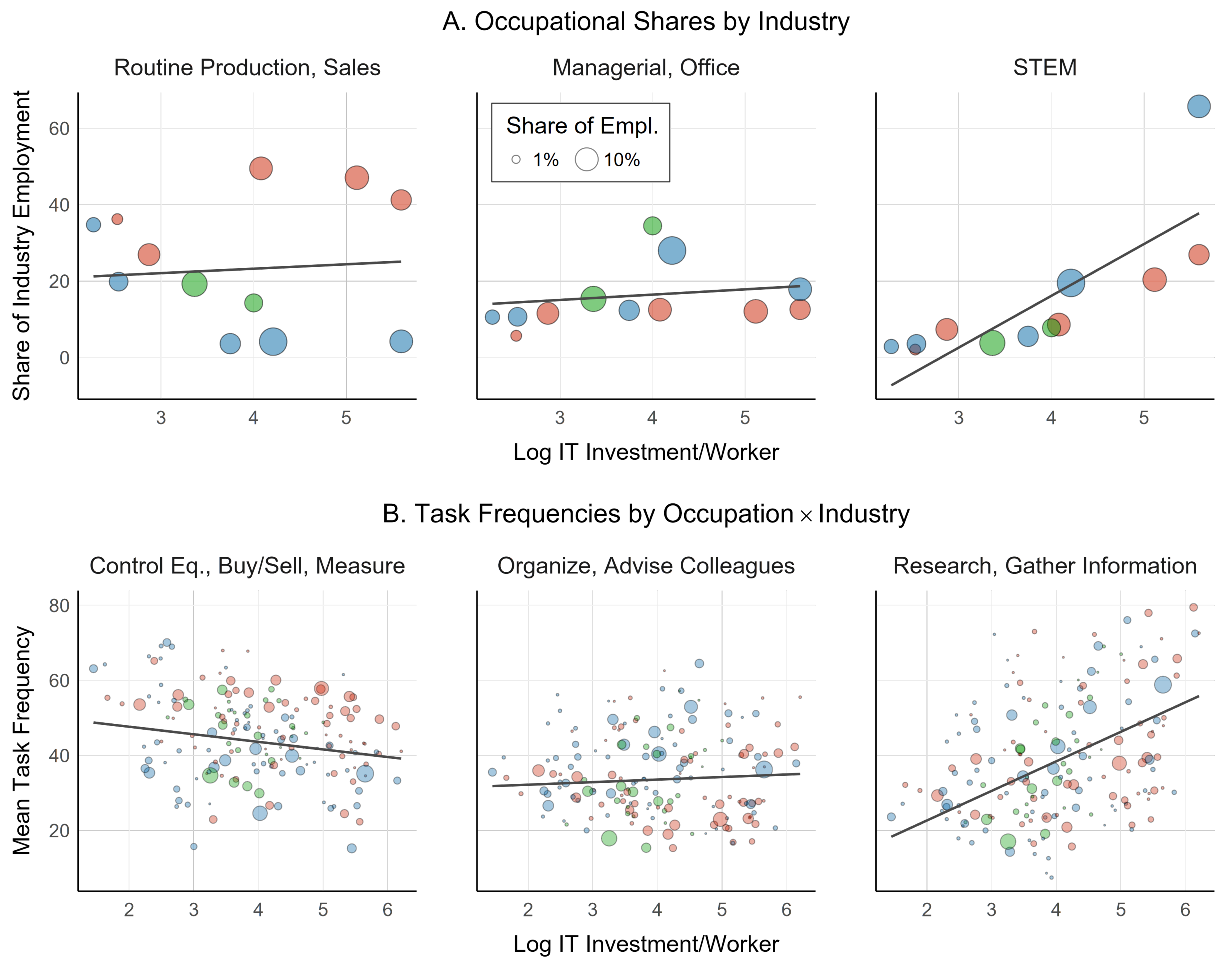

Figure 6.1: IT Investment and Labor Allocation, 2003-2010

Data: LIAB-AKM and 2006 BIBB Survey. Note: Weighted 2003-2010 employment shares and mean log ICT investment *per capita*. Panel A gives within-industry occupation shares for three groups: line workers, machinists, mechanics, and sales clerks; managers and office specialists; and scientists, engineers, and technicians. Panel B gives mean task frequencies across three sets of BIBB tasks: controlling machines, buying/selling, and measuring/testing; organizing work processes and advising colleagues; and research/design and gathering information. Line of best fit weighted by employment. Red indicates manufacturing industries, green commerce, and blue services.

Hence German wage sorting is associated with knowledge-intensity in production - an association consistent with a general mechanism based on sectoral production technology. This explanation would be convenient: it would straightforwardly account for the prevalence of wage sorting across the OECD, the importance of industry in the between-group results in Section 4, and the strengthening of wage sorting in recent decades. A more formal statement and test of the hypothesis lies outside the scope of this paper, however, and I leave it to future study.

6.3 Firm Pay and Experience

Finally, I turn to the relationships between worker experience, match tenure, and wage sorting, which appear to also be important given results in Tables 4.2 and 6.1. A role for match tenure is potentially problematic for this paper's conclusion that German wage sorting is technical in nature. If tenure predicts π and ϕ because sorted matches are less likely to separate, then this would instead suggest a mechanism based on assortative matching or sorting on unobservables, as argued previously in Section 4. On the other hand it could simply be that firms that pay a higher wage benefit from an older and more experienced workforce, as all workers are less likely to exit (*e.g.* due to less on-the-job search).[45]

I differentiate these two possibilities *via* a pair of regressions. The first attempts to rule out a selection mechanism:

$$\text{logit}(\text{Exit}_{p,t}) = \theta_0 + \theta_1 \pi_p + \theta_2 \phi_{f(p,t)} + \theta_3 \pi_p \times \phi_{f(p,t)} + \mathbb{I}_t + \mu_{p,t} \ , \tag{11}$$

where $\theta_2 < 0$ would indicate a general pay-tenure relationship, and $\theta_3 < 0$ a selection effect. The second regression considers whether π is increasing over time in experience and tenure:

$$\{w_{p,t}, \pi_{p,t}\} = \kappa_0 + \kappa_1 \text{Age}_{p,t} + \kappa_2 \text{Tenure}_{p,t} + \mathbb{I}_p + \tau_{p,t} \ . \tag{12}$$

Positive values of κ_1 and/or κ_2 would be indicative of a human capital mechanism, as opposed to experience/tenure capturing time-invariant, unobserved differences in ability or match quality. Regression (11) is performed on annual observations for the period 2003-2010, which corresponds to the approximate middle of the sample period, while (12) is estimated at the AKM panel level for the full sample period so as to capture within-person changes in π.[46]

Table 6.2 is consistent with a technical, human capital-based mechanism. Panel A shows that tenure is increasing in ϕ because *all* workers are less likely to exit high-paying firms, while in Panel B we see that π is increasing in tenure, largely due to a correlation with age. The coefficient on the match term is positive, indicating that sorted matches are more likely to exit, and not less. This could reflect nonlinearities in pay that the AKM regression fails to capture, or negative selection of sorted matches; but in either case it is inconsistent with a wage sorting mechanism based on assortative matching or sorting on unobservables. The negative coefficient on π, on the other hand, implies that any wage sorting - regardless of mechanism - will generate a positive correlation between ϕ and tenure. Hence tenure's

[45] In a directed search setting, a higher entry cost K will induce firms to pay a higher wage - one possible microfoundation for the assumption of rent-sharing in the conceptual framework.

[46] Note that panel dummies are not included in 12 as κ_1 would then be unidentified. Note as well that values of π and ϕ are demeaned, with $\pi \times \phi$ therefore greatest for high-high and low-low matches.

Table 6.2: Logit Regressions, Match Separation and Wage Effects, 2003-2010

	DV: Match Separation				DV: AKM Person Effect π		
	(1)	(2)	(3)	(4)	(1)	(2)	(3)
	A. Logit Regression, Match Separation and Wage Effects						
Firm Effect ϕ	**-1.44**		**-1.33**	**-1.33**			
	(.02)		(.02)	(.02)			
Worker Effect π		**-.88**	**-.77**	**-.67**			
		(.01)	(.01)	(.01)			
Interaction $\pi \times \phi$				**1.04**			
				(.07)			
	B. Panel Regression, Person Effect and Experience/Tenure						
Tenure					**.0070**		**.0021**
					(.0004)		(.0004)
Experience						**.0072**	**.0052**
						(.0004)	(.0004)
Person FE					X	X	X
Observations		7,618,822				10,920,183	
Persons						7,263,007	

DATA: LIAB-AKM. NOTE: Panel A: Logit regression of match separation indicator on 2003-2010 AKM wage effects. Separation is defined as worker exit from continuing establishments. Panel B: Panel regression of 1993-2017 AKM person effects on tenure and experience. Tenure is years in current establishment, and experience is years since first payment into social security. Bold text indicates statistical significance at the 5% level; robust standard errors shown.

explanatory power may, in part, be mechanical in nature.

These results cast additional light on the Gelbach decomposition in Table 6.1, and in particular on the portion of wage sorting associated with non-IT capital investment. Manufacturing sectors tend to be intensive in plant capital, and to employ an older, more experienced, but less-educated workforce. When comparing manufacturing and service industries with similarly-educated workers, we observe wage sorting occurring in parallel with physical capital and experience; whereas within a given sector, wage sorting is associated instead with IT investment and formal education. In the aggregate, IT and non-IT investment are positively correlated, but education and experience are negatively so. Hence these two sets of observable characteristics predict two distinct but offsetting vectors of wage sorting. This in turn explains why, in Table 4.2, neither education nor experience is by itself able to substantively account for the variation in π associated with $Cov(\pi, \phi)$.

7 Conclusion

Summarizing the main findings, West German wage sorting is (1) all or nearly all technical in nature, and predominantly a result of stable, *sector*-level wage differentials; (2) more prominent over time due to structural shifts associated with domestic outsourcing, and a general rise in skill premia; (3) strongly predicted by observable measures of physical and human capital, in particular IT and a formal STEM education; and (4) consistent with two general mechanisms, one related to industry production technology and the other to match separation rates. These results are consistent with our understanding of the major trends affecting labor markets, and intuitive given past work on sectoral wage gaps as in Krueger and Summers (1988) and Haskel and Slaughter (2002). They are also surprising. I find no evidence to indicate an assortative matching mechanism as suggested by Song et al. (2018), and more generally, my results are at odds with the current literature's focus on *match*-level heterogeneity in accounting for wage outcomes.

The role played by sectoral wage gaps has important policy implications. While German wage sorting is largely unrelated to collective bargaining coverage, much of its 1993-2017 increase is associated with a rise in temp agency employment. The equilibrium effects of this development on the wage distribution are unclear, and I leave them to a companion paper (Mouton, 2024), but the result directly suggests a relationship between wage sorting and the deregulation of the Hartz reforms. Similarly, while I find that wage sorting is associated with capital costs rather than implicit barriers to entry, this does not rule out a role for anti-competitive rents in output markets, if these pass through to wages. Policies that target sector-level competition may therefore also influence wage sorting, in addition to - and potentially interacting with - any other effects that they have.

It must be emphasized that the results in this paper are limited to observed wage sorting: if the AKM wage regression is mis-specified, or if bias causes within-sector wage sorting to go unobserved, then the findings presented here may be incomplete or themselves biased. In particular, I cannot reject assortative matching or any other mechanism *in general*, but only as an explanation for the empirical outcome I study. This limitation is important to note, and I have attempted to address the issue of limited-mobility bias to the extent possible. The motivation for this paper is, however, to attempt an explanation for the seminal findings of Card et al. (2013), and its methodological approach should be judged in that light.

An unanswered question is whether these results extend beyond Germany. Haltiwanger et al. (2024) link sectoral wage gaps to rising wage in equality in the U.S., and given the similarities with Germany previously documented by , it seems plausible that this paper's other findings will apply to the United States. More broadly, while the two mechanisms proposed in Section 6 are general in nature, a general prediction of wage sorting does not follow. First, these mechanisms depend on differences in sectoral capital-intensity, which may be region-specific. Second, they occur along separate dimensions of human capital - education and experience - which may be positively or, as in the German case, negatively correlated, with the overall effect depending upon relative magnitudes. These hypotheses do, however, generate predictions regarding observable characteristics: capital investment, education, and separation rates. Hence the question of generality is one that should, in principle, be straightforward to study.

References

Abowd, John M., Francis Kramarz, Paul Lengermann, Kevin L. McKinney, and Sébastien Roux (2012) "Persistent inter-industry wage differences: rent sharing and opportunity costs," *IZA Journal of Labor Economics*, 1 (1), 7.

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

Acemoglu, Daron and David Autor (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Card, David and Orley Ashenfelter eds. *Handbook of Labor Economics*, 4, 1043–1171: Elsevier.

Acemoglu, Daron and Pascual Restrepo (2018) "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108 (6), 1488–1542.

Akerman, Anders, Elhanan Helpman, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding (2013) "Sources of Wage Inequality," *American Economic Review*, 103 (3), 214–219.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser (2018) "Firms and the Decline in Earnings Inequality in Brazil," *American Economic Journal: Macroeconomics*, 10 (1), 149–189.

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*, 171 (3), 673–697.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen (2020) “The Fall of the Labor Share and the Rise of Superstar Firms.,” *Quarterly Journal of Economics*, 135 (2), 645–709.

Autor, David H. and David Dorn (2013) “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market,” *American Economic Review*, 103 (5), 1553–1597.

Autor, David H., Frank Levy, and Richard J. Murnane (2003) “The Skill Content of Recent Technological Change: An Empirical Exploration,” *The Quarterly Journal of Economics*, 118 (4), 1279–1333.

Bachmann, Ronald and Michael C. Burda (2010) “Sectoral Transformation, Turbulence and Labor Market Dynamics in Germany,” *German Economic Review*, 11 (1), 37–59.

Bagger, Jesper and Rasmus Lentz (2019) “An Empirical Model of Wage Dispersion with Sorting,” *Review of Economic Studies*, 86 (1), 153–190.

Bagger, Jesper, Kenneth Sørensen, and Rune Vejlin (2013) “Wage sorting trends,” *Economics Letters*, 118 (1), 63–67.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman (2016) “It’s Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States,” *Journal of Labor Economics*, 34 (S2), 67–97.

Bellman, Lisa, Ben Lochner, Stefan Seth, and Stefanie Wolter (2020) “AKM effects for German labour market data,” *FDZ-Methodenreport, IAB FDZ.*

Bonhomme, Stephane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler (2023) “How Much Should we Trust Estimates of Firm Effects and Worker Sorting?” *Journal of Labor Economics*, 41 (2).

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019) “A Distributional Framework for Matched Employer Employee Data,” *Econometrica*, 87 (3), 699–739.

Brown, Charles and James Medoff (1989) “The Employer Size-Wage Effect,” *Journal of Political Economy*, 97 (5), 1027–59.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson (2021) “The Productivity J-Curve: How Intangibles Complement General Purpose Technologies,” *American Economic Journal: Macroeconomics*, 13 (1), 333–372.

Buera, Francisco J., Joseph Kaboski, Richard Rogerson, and Juan I. Vizcaino (2022) "Skill-Biased Structural Change," *The Review of Economic Studies*, 89 (2), 592–625.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline (2018) "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36 (S1), S13–S70.

Card, David, Jörg Heining, and Patrick Kline (2013) "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*, 128 (3), 967–1015.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2022) "Matching in Cities.," *Journal of the European Economic Association*, 20 (4), 1478–1521.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2017) "Trade and Manufacturing Jobs in Germany," *American Economic Review*, 107 (5), 337–342.

Doms, Mark, Timothy Dunne, and Kenneth R. Troske (1997) "Workers, Wages, and Technology," *The Quarterly Journal of Economics*, 112 (1), 253–290.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth Troske (2004) "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment," *Journal of Labor Economics*, 22 (2), 397–430.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009) "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 124 (2), 843–881.

Engbom, Niklas and Christian Moser (2022) "Earnings Inequality and the Minimum Wage: Evidence from Brazil," *American Economic Review*, 112 (12), 3803–3847.

Fortin, Nicole M. and Thomas Lemieux (1997) "Institutional Changes and Rising Wage Inequality: Is There a Linkage?" *Journal of Economic Perspectives*, 11 (2), 75–96.

Franke, Reiner and Peter Kalmbach (2005) "Structural change in the manufacturing sector and its impact on business-related services: an input–output study for Germany," *Structural Change and Economic Dynamics*, 16 (4), 467–488.

Garicano, Luis (2000) "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108 (5), 874–904.

Garicano, Luis and Esteban Rossi-Hansberg (2015) "Knowledge-Based Hierarchies: Using Organizations to Understand the Economy," *Annual Review of Economics*, 7 (1), 1–30.

Garz, Marcel (2013) "Employment and wages in Germany since the 2004 deregulation of the temporary agency industry," *International Labour Review*, 152 (2), 307–326.

Gibbons, Robert and Lawrence Katz (1992) "Does Unmeasured Ability Explain Inter-Industry Wage Differentials," *The Review of Economic Studies*, 59 (3), 515–535.

Goldschmidt, Deborah and Johannes Schmieder (2017) "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure," *The Quarterly Journal of Economics*, 132 (3), 1165–1217.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii (2017) "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85 (1), 29–65.

Hall, Anja and Michael Tiemann (2021) "BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany 2006 (SUF)," *GESIS Data Archive, Cologne. ZA4820 Data file Version 1.0.1*, 22,0 MiB, 24,5 MB, Type: dataset.

Haltiwanger, John, Henry R. Hyatt, and James R. Spletzer (2024) "Rising Top, Falling Bottom: Industries and Rising Wage Inequality," *American Economic Review*, 114 (10), 3250–3283.

Haskel, Jonathan E. and Matthew J. Slaughter (2002) "Does the sector bias of skill-biased technical change explain changing skill premia?" *European Economic Review*, 46 (10), 1757–1783.

Håkanson, Christina, Erik Lindqvist, and Jonas Vlachos (2021) "Firms and Skills: The Evolution of Worker Sorting.," *Journal of Human Resources*, 56 (2), 512–538.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2020) "Leave-Out Estimation of Variance Components," *Econometrica*, 88 (5), 1859–1898.

Krueger, Alan B. and Lawrence H. Summers (1988) "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56 (2), 259–293.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante (2000) "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68 (5), 1029–1053.

Lentz, Rasmus (2010) "Sorting by search intensity," *Journal of Economic Theory*, 145 (4), 1436–1452.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2016) "Matching, sorting and wages," *Review of Economic Dynamics*, 19, 63–87.

de Melo, Rafael Lopes (2018) "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, 126 (1), 313–346.

Mouton, Andre (2024) "Skill-Bias, Firm-Bias, and Wage Inequality," *Working Papers*.

Schmidtlein, Lisa, Stefan Seth, and Matthias Umkehrer (2019) "Linked-Employer-Employee-Daten des IAB: Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993–2017," *FDZ-Datenreport, IAB FDZ.*

Shimer, Robert and Lones Smith (2000) "Assortative Matching and Search," *Econometrica*, 68 (2), 343–369.

Simón, Hipólito (2010) "International Differences in Wage Inequality: A New Glance with European Matched Employer–Employee Data," *British Journal of Industrial Relations*, 48 (2), 310–346.

Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter (2018) "Firming Up Inequality*," *The Quarterly Journal of Economics*, 134 (1), 1–50.

Spermann, Alexander (2011) "The New Role of Temporary Agency Work in Germany," *IZA Discussion Paper No. 6180*.

Spitz-Oener, Alexandra (2006) "Technical Change, Job Tasks and Rising Educational Demands: Looking Outside the Wage Structure," *Journal of Labor Economics*, 24, 235–270.

Tomaskovic-Devey, Donald, Anthony Rainey, Dustin Avent-Holt et al. (2020) "Rising between-workplace inequalities in high-income countries," *Proceedings of the National Academy of Sciences*, 117 (17), 9277–9283.

Torres, Sónia, Pedro Portugal, John T. Addison, and Paulo Guimarães (2018) "The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model," *Labour Economics*, 54, 47–60.

A Proofs of Main Results

A.1 Proof of Proposition 1

Defining the wage differentials

$$
\begin{aligned}
DF(z',\zeta',z,\zeta,s,\sigma) &= \log w(z',s,\zeta',\sigma) - \log w(z,s,\zeta,\sigma) \\
DW(s',\sigma',z,\zeta,s,\sigma) &= \log w(z,s',\zeta,\sigma') - \log w(z,s,\zeta,\sigma)\ ,
\end{aligned}
$$

and taking unconditional expectations over (z,s,ζ,σ), we obtain

$$
\begin{aligned}
&\mathbb{E}\big[DF(z',\zeta',z,\zeta,s,\sigma) \mid z',\zeta'\big] \\
&\quad = \sum_{s,\sigma}\left(\left[\sum_{s,\zeta} g(z,s) f_{z,s}(\zeta,\sigma)\right]\log w(z',s,\zeta',\sigma)\right) - \sum_{z,s,\zeta,\sigma} g(z,s) f_{z,s}(\zeta,\sigma)\log w(z,s,\zeta,\sigma) \\
&\quad = \log K(z') + \sum_{s,\sigma}\left(\left[\sum_{z,\zeta} g(z,s) f_{z,s}(\zeta,\sigma)\right]\log \hat{y}(z',s,\zeta',\sigma)\right) - Q\ ,
\end{aligned}
$$

and

$$
\mathbb{E}\big[DW(s',\sigma',z,\zeta,s,\sigma) \mid s',\sigma'\big] = \log C(s') + \sum_{z,\zeta}\left(\left[\sum_{s,\sigma} g(z,s) f_{z,s}(\zeta,\sigma)\right]\log \hat{y}(z,s',\zeta,\sigma')\right) - R\ ,
$$

where Q and R are scalars. Defining $F(s,\sigma) = \sum_z \left[g(z,s)\sum_\zeta f_{z,s}(\zeta,\sigma)\right]$ and likewise for $F(z,\zeta)$,

$$
\begin{aligned}
\mathbb{E}\big[DF(z',\zeta',z,\zeta,s,\sigma) \mid z',\zeta'\big] &= \log K(z') + \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z',s,\zeta',\sigma) - Q \\
\mathbb{E}\big[DW(s',\sigma',z,\zeta,s,\sigma) \mid s',\sigma'\big] &= \log C(s') + \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s',\zeta,\sigma') - R\ .
\end{aligned}
$$

Now we can write the difference between $\log w$ and $\mathbb{E}[\log w]$ as

$$
\begin{aligned}
&\log w(z',s',\zeta',\sigma') - \mathbb{E}[\log w(z,s,\zeta,\sigma)] = \\
&\qquad \mathbb{E}\big[DF(z',\zeta',z,\zeta,s,\sigma) \mid z',\zeta'\big] + \mathbb{E}\big[DW(s',\sigma',z,\zeta,s,\sigma) \mid s',\sigma'\big] \\
&\qquad + \sum_{z,s,\zeta,\sigma} g(z,s) f_{z,s}(\zeta,\sigma)\log\frac{\hat{y}(z',s'\zeta',\sigma')}{\hat{y}(z,s,\zeta,\sigma)} - \sum_{s,\sigma} F(s,\sigma)\log\frac{\hat{y}(z',s',\zeta',\sigma')}{\hat{y}(z',s,\zeta',\sigma)} \\
&\qquad - \sum_{z,\zeta} F(z,\zeta)\log\frac{\hat{y}(z',s',\zeta',\sigma')}{\hat{y}(z,s',\zeta,\sigma')}\ ,
\end{aligned}
$$

and substitution then gives us

$$
\log w(z',s',\zeta',\sigma') - \mathbb{E}[\log w(z,s,\zeta,\sigma)] =
$$

$$
\begin{aligned}
&\log K(z') + \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z',s,\zeta',\sigma) + \log C(s') + \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s',\zeta,\sigma') \\
&+ \log \hat{y}(z',s',\zeta',\sigma') - \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z',s,\zeta',\sigma) - \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s',\zeta,\sigma') \\
&- Q - R + S \ ,
\end{aligned}
$$

where S is a constant. Imposing the normalization $\mathbb{E}[\log w] = Q + R - S$ and defining

$$
\begin{aligned}
FE(z,\zeta) &= \log K(z) + \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z,s,\zeta,\sigma) \\
PE(s,\sigma) &= \log C(s) + \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s,\zeta,\sigma) \\
ME(z,s,\zeta,\sigma) &= \log \hat{y}(z,s,\zeta,\sigma) - \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z,s,\zeta,\sigma) - \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s,\zeta,\sigma) \ ,
\end{aligned}
$$

we have $\log w = FE + PE + ME$, with substitution for F yielding the result in the main text.

A.2 Proof of Proposition 2

I proceed by cases and evaluate the implications of each wage sorting mechanism for the covariance of PE and FE within-(z,s), between-z, and between-s.

Technical sorting. Suppose that $Cov\Big(\log K(z), \log C(s)\Big) > 0$. Because K is constant given z, and C constant given s, it follows trivially that the covariance conditional on z, s, and (z,s) will be zero. From the law of total covariance it must be that both between-group covariances (z and s) are positive, and the result is shown.

Positive assortative matching. That positive assortative matching results in a positive covariance within-(z,s) is trivial, given the assumptions that $\frac{f_{z,s}(\zeta_+,\sigma_+)}{f_{z,s}(\zeta_+,\sigma_-)} > \frac{f_{z,s}(\zeta_-,\sigma_+)}{f_{z,s}(\zeta_-,\sigma_-)}$ and y is monotonically increasing. That it will be absent between-z, -s, and -(z,s) follows from the fact that higher values of $\sum f_{z,s}(\zeta,\sigma)y(\zeta,\eta)$ do not affect $\mathbb{E}[\hat{y}(z,s,\zeta,\sigma) \mid z,s]$, and therefore cannot affect $\mathbb{E}[w(z,s,\zeta,\sigma) \mid z,s]$. If market-level variation in $f_{z,s}$ raises the expected value of either PE or FE, then the expected value of the other must be lower.

I show this result explicitly for the simplified case in which $f_{z_H,s_H}(\zeta_+,\sigma_+) = 1$, $f_{z_L,s_L}(\zeta_-,\sigma_-) = 1$, and $g(z_H,s_H) + g(z_L,s_L) = 1$. Under these assumptions we will have

$$
\begin{aligned}
\hat{y}(z_H,s_H,\zeta,\sigma) &= \frac{y(\zeta,\sigma)}{y(\zeta_H,\sigma_H)} \\
\hat{y}(z_L,s_L,\zeta,\sigma) &= \frac{y(\zeta,\sigma)}{y(\zeta_L,\sigma_L)} \ .
\end{aligned}
$$

Defining $F(s,\sigma)$ and $F(z,\zeta)$ as in the previous proof, it is straightforward to show that

$$\sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z,s,\zeta,\sigma) = \sum_{s,\sigma} F(s,\sigma)\log \hat{y}(z',s,\zeta',\sigma)$$
$$\sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s,\zeta,\sigma) = \sum_{z,\zeta} F(z,\zeta)\log \hat{y}(z,s',\zeta,\sigma') \ ,$$

for any (z,z',s,s'). But if the expected values of PE and FE are constant across markets, the it follows directly that $Cov\Big(\sum_{s,\sigma} F(s,\sigma)\log\hat{y}(z,s,\zeta,\sigma), \sum_{z,\zeta} F(z,\zeta)\log\hat{y}(z,s,\zeta,\sigma)\Big) = 0$ between-s, between-z, and between-(z,s).

Unobserved sorting. Focusing on the case in which sorting occurs on ability (3a), that $K(z)$ is constant conditional on z implies that $Cov\Big(\log K(z), \sum_{z,\zeta} F(z,\zeta)\log\hat{y}(z,s,\zeta,\sigma)\Big) = 0$ within-z; hence the between-z covariance must be positive and the between-s covariance must equal zero. On the other hand the term $\sum_{z,\zeta} F(z,\zeta)\log\hat{y}(z,s,\zeta,\sigma)$ is a market-level aggregate and is, by construction, a constant conditional on (z,s), and for this reason the covariance will be zero conditional on (i,j). The same steps show the equivalent result for the case in which sorting occurs on productivity.

A.3 Proof of Proposition 3

As the result is unaffected by the inclusion of time-varying regressors, I consider a simplified regression equation of the form

$$w_{p,t} = \pi_p + \phi_{f(p,t)} + \epsilon_{p,t} \ .$$

Letting hat variables indicate OLS estimates, the property that residuals sum to zero conditional on p and f implies that

$$\hat{\pi}_p - \pi_p = \frac{1}{N_p}\sum_t \Big[\epsilon_{p,t} - \big(\hat{\phi}_{f(p,t)} - \phi_{f(p,t)}\big)\Big]$$
$$\hat{\phi}_f - \phi_f = \frac{1}{N_f}\sum_{p,t} \Big[\mathbb{I}(f(p,t)=f)\Big]\Big[\epsilon_{p,t} - \big(\hat{\pi}_p - \pi_p\big)\Big] \ ,$$

where N_p and $N_{f(p,t)}$ are the number of observations associated with individual p and firm f, respectively. Define $\sigma_{f(p,t)} = \hat{\phi}_{f(p,t)} - \phi_{f(p,t)}$ to be the estimation error in firm f's fixed effect, and note that for workers who are only observed at firm f', it must be that $\frac{1}{N_p}\sum_t \big(\hat{\phi}_{f(p,t)} - \phi_{f(p,t)}\big) = \sigma_{f'}$. Hence we can rewrite the previous equations as:

$$\hat{\pi}_p - \pi_p = \frac{1}{N_p}\sum_t \epsilon_{p,t} - \frac{1}{N_p}\sum_t \sigma_{f(p,t)} \tag{13}$$

$$\sigma_f = \frac{1}{N_f}\sum_{k,u}\left[\mathbb{I}\big(f(k,u)=f\big)\Big(\epsilon_{k,u} - \frac{1}{N_k}\sum_t \epsilon_{k,t} + \frac{1}{N_k}\sum_t \sigma_{f(k,t)}\Big)\right]$$
$$= \frac{1}{N_f^m}\sum_{k,u}\left[\mathbb{I}\big(\text{mover}, f(k,u)=f\big)\Big(\epsilon_{k,u} - \frac{1}{N_k}\sum_t \epsilon_{k,t} + \frac{1}{N_k}\sum_t \sigma_{f(k,t)}\Big)\right] \ , \qquad (14)$$

where N_f^m is the number of observations of firm f for which a given worker is a job-mover, and observed at multiple firms within the AKM connected set. If the number of periods in the sample is T, then the variance of $\frac{1}{N_p}\sum_t \epsilon_{p,t}$ is at least σ_ϵ^2/T (*i.e.* given that we will have T or fewer observations of p), and it follows that the variance of $\hat{\pi}_p - \pi_p + \frac{1}{N_p}\sum_t \sigma_{f(p,t)}$ will also be at least σ_ϵ^2/T. On the other hand the variance of the right-hand side of (14) will be at least as large as $\frac{1}{4}\frac{\sigma_\epsilon^2}{N_f^m}$. Hence it must be that the variance of σ_f is non-zero, in which case the covariance term

$$Cov\big(\hat{\pi}_p, \hat{\phi}_{f(p,t)}\big) = Cov\left(\pi_p + \frac{1}{N_p}\sum_t \epsilon_{p,t} - \frac{1}{N_p}\sum_t \sigma_{f(p,t)},\ \phi_{f(p,t)} + \sigma_{f(p,t)}\right) \ , \qquad (15)$$

will be downward-biased.

Now let $g \in G$ define a partition of the sample and suppose that the true values of π and ϕ consist of a group component and an idiosyncratic term,

$$\pi_p = \pi_{g(p)} + \mu_p$$
$$\phi_f = \phi_{g(f)} + \rho_f \ ,$$

where $\mathbb{E}[\mu] = \mathbb{E}[\rho] = 0$. For the sake of exposition, let $g(p)$ and $g(f)$ be time-invariant. Then

$$\hat{\pi}_p = \pi_{g(p)} + \mu_p + \frac{1}{N_p}\sum_t \epsilon_{p,t} - \frac{1}{N_p}\sum_t \sigma_{f(p,t)}$$
$$\hat{\phi}_f = \phi_{g(f)} + \rho_f + \sigma_f \ ,$$

and averaging across group g we obtain

$$\hat{\pi}_g = \sum_{p\in g}\frac{N_{p,g}}{N_g}\pi_g + \sum_{p\in g}\sum_{p\in g}\frac{N_{p,g}}{N_g}\frac{1}{N_p}\sum_t \epsilon_{p,t} - \sum_{p\in g}\frac{N_{p,g}}{N_g}\frac{1}{N_p}\sum_t \sigma_{f(p,t)}$$
$$= \sum_{p\in g}\frac{N_{p,g}}{N_g}\pi_g + \sum_{p\in g}\sum_{p\in g}\frac{1}{N_g}\sum_t \epsilon_{p,t} - \sum_{p\in g}\frac{1}{N_g}\sum_t \sigma_{f(p,t)}$$
$$\hat{\phi}_g = \sum_{f\in g}\frac{N_{f,g}}{N_g}\phi_g + \sum_{f\in g}\frac{N_{f,g}}{N_g}\sigma_f \ ,$$

where N_g and $N_{i,g}$ give the total and person-/firm-conditional number of observations in group g. As N approaches infinity, N_g will do likewise; and provided that $N_{p,g}/N_g$ and $N_{f,g}/N_g$ therefore converge to zero (*i.e.* the number of observations per person or firm does not increase), we will have $\sum_{p \in g} \frac{1}{N^g} \sum_t \epsilon_{p,t}$ and $\sum_{f \in g} \frac{N_{f,g}}{N^g} \sigma_f$ converging to zero. In this case estimates of π^g and ϕ^g will be consistent and the covariance term

$$Cov\big(\hat{\pi}_g, \hat{\phi}_g\big) = Cov\left(\sum_{p \in g} \frac{N_{p,g}}{N_g} \hat{\pi}_{g(p)} + \sum_{p \in g} \frac{1}{N^g} \sum_t \epsilon_{p,t} - \sum_{f \in g} \frac{N_{f,g}}{N^g} \sigma_f,\ \sum_{f \in g} \frac{N_{f,g}}{N_g} \hat{\phi}_{g(f)} + \sum_{f \in g} \frac{N_{f,g}}{N^g} \sigma_f\right) ,$$

will be unbiased as $N \to \infty$.

Note that the proof is not substantively changed by allowing groups to vary across time or within firms, provided that the key assumptions (convergence to zero of $N_{p,g}/N_g$ and $N_{f,g}/N_g$ as N_g becomes large) continue to hold.