The Role of Firms and Occupations in Wage Inequality

Luke Heath Milsom* Shihang Hou*

February 6, 2025

Abstract

This paper quantifies the relative importance of firms and occupations to wage inequality. We estimate a worker-job two-way fixed effect model using French data, allowing for unrestricted firm-occupation interactions, and use our model to decompose job premia variance and worker-job sorting to between- and within- occupation components. We find that between-occupation variance in pay premia is as large as within-occupation variance and sorting of workers between occupations is substantially more important than sorting of workers between firms within occupations. Our results suggest that measurements of worker-firm sorting which ignore occupations may conflate sorting to firms with sorting to occupations.

*Heath Milsom: Institute for Fiscal Studies and KU Leuven, luke_m@ifs.org.uk. Hou: Institute for Employment Research (IAB), shihang.hou@iab.de. With thanks to Abi Adams, Stephane Bonhomme, Wolfgang Dauth, Binta Zahra Diop, Lucas Finamor, Martin Friedrich, Simon Janssen, Vatsal Khandelwal, Adrian Lerche, Sanghamitra Mukherjee, Barbara Petrongolo, Simon Quinn, Malte Sandner, Margaret Stevens, Martin Weidner, Verena Wiedemann, and Hannah Zillessen for helpful feedback and comments in the course of writing this paper. Access to some confidential data, on which this work is based, was made possible within a secure environment provided by CASD – Centre d'accès sécurisé aux données (Ref. 10.34724/CASD).

How much you earn depends on where you work and what you do for work. Consequently, heterogeneity in firm and occupation wage premia and the propensity of high-wage workers to sort into high-wage firms or high-wage occupations are important components of aggregate wage inequality. The role of firms and the role of occupations are well documented in isolation,[1] but few papers have looked at both factors in conjunction. Designing effective policies to address wage inequality requires quantifying the relative importance of each factor. If high-paying occupations are often concentrated in high-wage firms, worker-firm sorting could reflect the sorting of workers to occupations, and firm premia heterogeneity the clustering of occupations to firms. For example, high skilled workers may sort to Google because they hire a lot of skilled software engineers, a high-paying occupation, not because Google is a high-wage employer within those occupations. Not addressing this conflation could lead to ineffective policy recommendations which over emphasize the role of either firms or occupations.

In this paper, we study the interaction of firms and occupations in wage determination and quantify the degree of heterogeneity and worker sorting along each dimension. First, we augment the traditional worker-firm two-way fixed effect model of Abowd et al. [1999] (AKM) by estimating job fixed effects instead of firm fixed effects, where a job is a firm-occupation pair, e.g. software developer at Google, or industrial systems engineer at BMW.[2] The interacted fixed effect allows us to then decompose the variance of job fixed effects and the covariance between worker and job fixed effects, into a between-occupation component and a within-occupation between-firm component. By allowing for completely flexible interactions between firms and occupations we can study the relative contribution of each element controlling for possible clustering of certain occupations in certain firms. Finally, we use this decomposition to distinguish between wage inequality due to the kind of firm wage premia heterogeneity in the AKM literature and inequality due to variation in occupation

[1] Autor et al. [2003], Goos and Manning [2007], Acemoglu and Autor [2011] and other papers suggest that wages depend on the tasks performed on the job (e.g. manual, interactive, and abstract tasks), and argue that technological changes like automation and computerisation have differentially affected the price of different tasks, leading to polarisation of job earnings depending on the tasks the jobs involve. Abowd et al. [1999], Mortensen [2003], Card et al. [2013, 2016] and Song et al. [2019] have documented that firms pay similar workers differently; recent estimates from different countries [Bonhomme et al., 2023] suggest that these differences accounting for 6-16% of total log wage variance.

[2] This is equivalent to interacting the firm fixed effect in AKM models with occupation fixed effects.

wage premia studied in the task and occupation literature.

We apply the worker-jobs model to a panel dataset covering the universe of French workers [Babet et al., 2022], correcting for bias due to limited mobility using the leave-one-out approach due to Kline et al. [2020][3]. We show evidence for exogenous conditional moves across jobs using event studies [Card et al., 2013] and robustness of our results to using only moves between firms, which are standard in worker-firm AKM models. Finally, we show external validity by confirming our main findings in a comparable administrative matched employee-employer dataset for Germany.

We draw two primary conclusions from our analysis. First, we find that adopting a decomposition based on a worker-jobs model reduces the share of variance attributed to individual heterogeneity and increases the share attributed to job heterogeneity and worker-job sorting. We find that 49% of wage variance can be explained by individual heterogeneity, 11% by job heterogeneity, and 21% by the sorting of higher-wage workers to higher-wage jobs. Relative to the worker-firm AKM model, decomposing our worker-job model increases the share due to job FE by 4.2 pp. (65.6%) and the share due to sorting by 11.9 pp. (130.8%), and decreases the share due to individual FE by 8.1 pp (14.1%). The differences between the decompositions can be explained by noting that if the worker-job model were the true model, then estimating the worker-firm model would lead to an omitted variable — the contribution of the worker's occupation to his wage relative to the mean wage at his firm. The variance of this omitted variable is distributed in log wage decompositions depending on the covariance between this omitted variable and the individual fixed effect (reflecting worker-occupation sorting), or the firm fixed effect (reflecting sorting of particular occupations to particular firms).

Second, we show that (i) the variance of job fixed effects can be broken down equally into variation between occupations, and variation within occupations between firms, but (ii) most of the observed covariance between individual and job fixed effects is due to sorting between workers and occupations and not due to sorting between workers and firms within occupations. Decomposing the covariance between worker- and job- fixed effects using the law

[3]Occupations are measured at the 4-digit level which gives 430 unique occupations. Examples of occupations are: midwives, professors and lecturers, surveyors and topographers, and qualified roofers.

of total covariance, we findthat the share of log-wage variance due to sorting of individuals to occupations (17.2%) if four times as large as the share due to sorting of workers between firms within occupations (4.0%). In contrast, we find that 48% of the variance of job fixed effects (4.8% of the total residualised log wage variance) is due to between-occupation variation and 52% (5.2% of the total residualised log wage variance) is due to within-occupation variation, a fairly even split. Our results stand in contrast to the conclusions from a decomposition of the worker-firm fixed effects model which finds that 9.1% of residualised log wage variance is due to worker-firm sorting. We further investigate whether the importance of sorting could depend on whether the occupation was highly paid; to do so, we decompose the variance of residualised log wages conditional on occupation into components due to worker heterogeneity, firm heterogeneity, and sorting between workers and firms within occupations. We find that for high-pay occupations, individual and firm heterogeneity are both relatively more important, but we do not find evidence that there is greater sorting of workers to higher-paying firms within these occupations.

By flexibly including firms and occupations in our specification we can estimate firm-occupation match effects in wage determination and worker-job sorting. This effects could be important, after all, Google may not pay its canteen chefs the same premium as its software engineers since software is a central part of their business, and as a result, the best engineers may go to Google, but the best chefs may work at Michelin-star restaurants. We find substantial heterogeneity in occupation wage premia within firms, and in particular, high-wage occupations are typically highly paid regardless of the firm that they work in. Our results cohere with models where firms may offer different wage premia to workers in different occupations in ways that cannot be captured by a log-additive occupation fixed effect. For example Bloesch et al. [2022] suggest that occupations in some sense more central to the firm have greater hold-up power in pay negotiation.

Finally, we use our framework to analyse whether there is heterogeneity in the determinants of wage inequality across commuting zones. We find that in larger commuting zones, individual and firm heterogeneity explain a larger share of total residualised log-wage variance, but as in the case of more highly paid occupations we do not find evidence that there is greater sorting of workers to higher-paying firms within occupations. Larger commuting zones instead have more assortative sorting of workers between occupations than smaller ones. These results cohere with the finding in the aggregate decompositions that there seems to be fairly little within-occupation between-firm sorting despite there being substantial within-occupation between-firm pay premia.

Our work contributes to the large and active literature exploring the factors underlying wage inequality using the two-way fixed effects approach. In a variety of countries over many time periods, papers have documented that firms pay different wage premiums to similar workers, and have quantified the magnitude of this firm wage dispersion as a share of overall wage inequality Abowd et al. [1999, 2002], Card et al. [2013], Song et al. [2019], Card et al. [2018], Babet et al. [2022].[4] We contribute to this literature by analysing a worker-jobs fixed effect model which allows us to distinguish between the roles of occupations and firms in wage determination. We emphasise three conclusions in contrast to the existing literature. First, we document that with a worker-job model, the share of variance attributed to individual fixed effects decreases and the share attributed to job fixed effects and sorting increases. We argue that this is consistent with there being an omitted variable in the worker-firm model when the worker-job model is the true model. Second, our approach finds relatively little sorting between workers and firms conditional on their occupation and instead finds that most sorting that we measure seems to be sorting between workers and occupations. On the other hand, we find a substantial role for between-occupation pay dispersion which we find to be quantitatively about as important as within-occupation, between-firm wage heterogeneity. These two conclusions coheres with the focus on analysing task content and skill prices as sources of inequality (e.g. Autor et al. [2003], Goos and Manning [2007], Acemoglu and Autor [2011]); our work contributes to this literature by providing a measure

[4] Card et al. [2013] also provides, in section VII of their paper, an analysis of changes in between-occupation wage differentials using the AKM model, concluding that the largest factor, accounting for 42% of such changes, was "a rise in the covariance between the mean person effect in an occupation and the mean establishment wage premium for that occupation". They interpret this as "people in higher-paid occupations [being] increasingly concentrated at establishments that pay all workers a higher wage premium". However, we argue that the finding of a substantial covariance between within-occupation mean individual fixed effect and within-occupation mean establishment fixed effect across occupations could also be explained in a world without establishment wage premia and only occupation wage premia. In that case, we might expect there to be a similar covariance simply because high-premia occupations are concentrated in certain firms. We might also get the kind of increases documented in Card et al. [2013] even without any heterogeneity in firm fixed effects if workers in high-wage occupations are increasingly concentrated within particular firms.

of the relative importance of these factors in overall wage inequality. Third, although we find greater dispersion in pay premia between firms within occupations in larger occupations and larger labour markets, we do not find evidence that this greater dispersion leads to more sorting between firms, as e.g. Dauth et al. [2022] finds in the case of sorting across local labour markets.

A few other papers have also explored extensions to AKM where firms could pay different wages to workers in different occupations. The paper closest in spirit to ours is Torres et al. [2018], which analyses a model with log-additive worker, firm and occupation fixed effects. Relative to their approach, we do not require firm and occupation fixed effects to be log-additive and in fact show that the log-additive specification does not fit the data especially for high-wage occupations in high-wage firms. Goldschmidt and Schmieder [2017] and Lamadon et al. [2022] both interact the firm fixed effect with broad occupation categories although neither paper focuses on distinguishing between firms and occupations in decomposing wage inequality. Relative to their work which restricts firm wage premia to vary only across broad occupation categories, we allow for unrestricted interaction of occupations with firms.

Our work also contributes more generally to the literature on wage determination by documenting substantial degrees of within-firm heterogeneity in pay premia between occupations. This departs from many current models, that is concerned with either firms or occupations and abstracts completely from the other. Our work provides empirical support for some recent models (e.g. Bloesch et al. [2022], Haanwinckel [2023]) which allow for wage premia to vary across occupations within firms.

The rest of our paper proceeds as follows. Section 1 sets out the formal econometric model we study in this paper. Section 2 describes the data used in this study. We present our results in two sections: section 3 presents our main decomposition results, summarises our specification tests and robustness checks (subsection 3.1), and presents our further decomposition of job fixed effect variance and the covariance between workers and job fixed effects into between- and within-occupation components (subsection 3.2). We also present our analysis of the variation of wage premia within a firm in subsection 3.3. Section 4 considers whether log wage decompositions vary between occupations and commuting zones. Finally, section 5 concludes.

1 Worker-job two way fixed effects econometric model

In our econometric framework, we decompose log wages into components due to worker specific heterogeneity, job specific heterogeneity, and to the sorting of workers into jobs. For each individual $i \in \mathcal{I}$ in each period $t \in \mathcal{T}$ we observe the firm $F(i,t) \in \mathcal{F}$ and occupation $O(i,t) \in \mathcal{O}$ they work in as well as the wage w_{it} received. We denote by $J(i,t)$ the *job* an individual i is employed doing, in period t where j is a firm-occupation pair i.e. $J(i,t) \in \mathcal{J} = \mathcal{F} \times \mathcal{O}$. We assume that log wages $\ln(w_{it})$ can be described by the sum of worker, α_i, and job, $\lambda_{J(i,t)}$, components as well as an idiosyncratic error as shown in equation 1. Throughout we additionally condition on a set of time varying worker covariates X_{it} such as age and year fixed effects. These have been omitted in this discussion for notational brevity. In practice, for computational reasons, we follow the recommendation in Kline et al. [2020] to apply our models to $ln(w_{it}) - X_{it}\hat{\beta}$ throughout, where X_{it} comprises of a cubic age profile and year fixed effects.

$$\ln(w_{it}) = \alpha_i + \lambda_{J(i,t)} + \varepsilon_{it} \tag{1}$$

This setup follows the literature started by Abowd et al. [1999] and continued by others such as Card et al. [2013] and Song et al. [2019]. In this model, person fixed effects α_i can be interpreted as a combination of skills and other factors that are awarded equally across jobs j. Similarly, λ_j can be interpreted as a pay premium offered to all workers in job j irrespective of unobserved worker-specific characteristics. Note that the usual worker-firm models (as opposed to worker-job) use ψ_f instead of λ_j and interpret it as a pay premium offered to all workers at a specific firm. This means that, for example, software developers, administrators, managers, cleaners, and chefs at Google all receive the same log-additive pay premium.

Our model nests this as a special case where all occupations within a firm have the same λ_j, but in general is significantly more flexible by allowing this premium to vary within a firm by occupation. For example, this would allow Google to pay its software developers a higher premium than its chefs and the Ritz to pay a higher premium to its chefs than its software developers.

The fixed effects in the worker-firm model are identified by *movers* between firms. In the same way, the job fixed effects in our model are identified by the same individual moving from one job to another and noting how their wages changed due to the move, and individual fixed effects were identified by observing workers' wages between jobs relative to other workers in those jobs.

Denote by y the stacked vector of individual wages, $A = [a_1, \cdots]$ the design matrix of individual indicators, $J = [b_1, \cdots]$ the design matrix of job indicators and ε the stacked vector of error terms. Then we can write equation 1 as $y = A\alpha + J\lambda + \varepsilon$ or more succinctly as $y = Z\beta + \varepsilon$ where $Z = [A, J]$ and $\beta = [\alpha, \lambda]$. We will estimate equation 1 by OLS to find $\hat{\beta} = \beta + (Z'Z)^{-1}Z'\varepsilon$. This result highlights the two main hurdles facing identification. First note that implicit in the above derivation is the invertibility of $Z'Z$. This is not the usual innocuous assumption; in general $Z'Z$ is not invertible in this set up if there are *unconnected* individuals or firms. That is, $Z'Z$ is only invertible in the subset of the data representing the largest connected set; since the set of most connected jobs is likely to be selected, our results may not be representative of the full population of jobs. Second, the estimates of the worker and job fixed effects are only unbiased under the standard OLS assumption of exogeneity. We discuss this assumption in more detail in subsection 3.1.

The specification in equation 1 allows a simple decomposition of the observed variance of log wages as given in equation 2 below.

$$\mathbb{V}[\ln(w_{ijt})] = \underbrace{\mathbb{V}[\alpha_i]}_{\substack{\text{Variance due to individual} \\ \text{heterogeneity}}} + \underbrace{\mathbb{V}[\lambda_j]}_{\substack{\text{Variance due to job} \\ \text{heterogeneity}}} + \underbrace{2 \cdot Cov[\alpha_i, \lambda_j]}_{\substack{\text{Variance due to workers} \\ \text{sorting into jobs}}} + v \tag{2}$$

This shows how, for each period, wage variance can be decomposed into four components. $\mathbb{V}[\alpha_i]$ captures the variance due to individual characteristics, and $\mathbb{V}[\lambda_j]$ similarly captures the variance due to job characteristics. Finally, $2Cov(\alpha_i, \lambda_j)$ captures the component of the variance which is due to sorting or matching of individuals to jobs, and v captures unexplained variance (as well as variances and covariances relating to included control variables).

1.1 Limited mobility bias

Limited mobility bias refers to the finding (Andrews et al. [2008], Andrews et al. [2012]) that estimates of the second (or higher) order moments of fixed effects based on fixed effect estimates are typically biased, even though the estimates are not biased themselves. This bias is a form of the incidental parameter problem and arises from the fixed effects being estimated with large standard errors when panels are short and/or when there is limited mobility of workers across firms. Jochmans and Weidner [2019] show how to understand the likely size of the bias as a function of connectedness of the network of workers and firms. Bonhomme et al. [2023] analyse the size of the bias based on simulated data as well as actual data from the US as well as Austria, Norway, Italy, Sweden, and find that the bias is both empirically important and can qualitatively affect conclusions about sorting and inequality. As discussed in Andrews et al. [2008] and Bonhomme et al. [2023] the bias intuitively arises due to a lack of movers. Job fixed effects are estimated from movers, and so few movers per-job means that they will be estimated with more error (although estimates will remain unbiased). More variation in job fixed effects estimates implies a greater estimated variance in such fixed effects. Then, as worker effects and job effects enter in a linearly additive manner on average higher estimated job fixed effects will be associated with lower estimated worker effects and therefore the covariance between job and worker effects will be underestimated. In our setting by considering jobs rather than firms we increase the number of movers by 94%. However, there are 3.5 times as many unique jobs as there are firms, and so the number of movers per firm/job decreases from 2.9 to 1.6 in our data.

In our main analysis we adopt the leave-one-out bias correction approach considered by Kline et al. [2020], and show the robustness of our main results to the homoskedastic version of this correction due to Andrews et al. [2008]. Andrews et al. [2008] derives a formula for limited mobility bias for variance components and propose a correction for the bias based on the assumption of homoscedasticity. Kline et al. [2020] relaxes the assumption of homoscedasticity and derives bias corrections for when heteroscedasticity is present.[5] They

[5]We compute both corrections using the Pytwoway package (Bonhomme et al. [2023]) Following the recommendation of Kline et al. [2020], we first residualise the wage variable from the age profile and year fixed effects, before running the fixed effect estimator without controls. Due to the large computational cost of computing the KSS adjustments exactly, this package approximates bias adjustment using the Johnson-

do this by providing an unbiased and consistent estimator for individual variances σ_{it}^2 based on leave-out coefficient estimates. Since the heteroscedasticity correction is estimated using leave-one-out estimates of the individual-specific variance of the error ε_{it}, it can only be estimated on a leave-one-out connected set, that is, a set of workers and firms that remain connected when any one observation is removed from the sample.

A downside of this approach is that the leave-one-connected set can be substantially smaller than the connected set, and expose the estimates to sample-selection bias. Larger firms/ jobs are more likely to belong to the connected set, and thus to the extent to which we expect differences along the firm/ job size distribution one should consider our results as mainly pertaining to larger firms/jobs. In our data section, we show that this is a possible concern in our study. For this reason we show that the similar conclusions can be reached using the homoskedastic correction due to Andrews et al. [2008] which allows us to use the same estimand as in a naive AKM decomposition in appendix section A.2.2.

1.2 Identification of fixed effects

To uncover unbiased estimates of the fixed effects, we require that $\mathbb{E}[Z'\varepsilon|Z] = 0$, the usual exogeneity assumption in OLS. The validity of this assumption for the classic worker-firm framework is discussed in detail in Card et al. [2013], and in our paper, we focus on what changes in our context relative to the standard AKM model. As identification comes from movers across jobs, a sufficient condition is: $\forall j, Pr(J(i,t) = j|\varepsilon) = Pr(J(i,t) = j)$; this is therefore what we focus our attention on.

There are two ways in which the assumption could be violated (Card et al. [2013]). First, there might be match effects that are not captured by worker and job fixed effects, i.e. if workers sort to jobs on the basis of a worker-job match-specific characteristic not captured by α_i and $\lambda_{J(i,t)}$. Relative to the AKM specification, this is more of a concern, as the skills necessary for different occupations are unlikely to be unidimensional as implied in our model (see e.g. Lindenlaub [2017] and Lise and Postel-Vinay [2020] for work arguing for the importance of multi-dimensional skills). Thus, there may be dimensions of individual skill not captured by the fixed effect that generate match-specific surplus at particular occupations

Lindenstrauss approximation.

which are pushed to the error term. On the other hand, the assumption coheres with the Acemoglu and Autor [2011] task model since they make the assumption that tasks can be sorted in terms of complexity, where more skilled workers have a comparative advantage in performing more complex tasks.

A second possible problem is that temporary variation in wages may be correlated with the job that workers perform. A concern of this type in Card et al. [2013] is that the statistical model is incompatible with models of the labour market where workers move to jobs due to high transitory wage offers not related to the firm fixed effect.[6] In our model, an additional problem of this kind is if temporary occupation-specific productivity shocks lead to substantial movement between occupations in the period.

If the econometric model specified is correct, then when a worker moves from a job with a high wage premium to a job with a low wage premium, we should see a step-wise relative change in their earnings that is roughly equal to the negative of an analogous move from the low-wage premium job to the high wage premium job. On the other hand, if there were match effects or temporary wage effects not captured by the fixed effects, then we should expect such moves between jobs to not necessarily lead to a relatively symmetric effect on earnings. Card et al. [2013] produce a diagnostic for the identification assumption based on this idea as follows: they first cluster firms into four clusters by their wage premia, and then study wage changes when workers move between firms in these clusters. They argue that if the identification assumption holds, moves between firm clusters should produce relatively symmetric wage changes. We use similar event study diagrams to show that wage moves between firm and job fixed effect quartiles are symmetric. These exercises are reported in section 3.1.1.

Finally, another possible criticism of this approach is that the two-way fixed effect regression imposes an inappropriate log-additive functional form on worker and job fixed effects. The log-additive structure we impose on the data could mask important heterogeneity, mechanisms or disallow potentially relevant theoretical channels. To check for this we follow the approach in Card et al. [2013] and show that the mean of the residuals of the two-way FE regression are near zero on average, and by job and worker cells. This exercise and its results

[6] A prominent model with such a mechanism is Postel-Vinay and Robin [2002].

are summarised in appendix A.5. We find that the specification performs well in this test.

1.3 Interpretation of fixed effects

Individual fixed effects can be interpreted as the component of log-wages that are individual specific and time-invariant conditional on one's job. How does the interpretation of the individual fixed effects in our model differ from the interpretation of the individual fixed effects in the standard framework? The key difference lies in the interpretation of the role of occupations in wage determination. In the standard AKM setting, a high individual fixed effect could reflect that a worker tends to work in high-productivity occupations, possibly because they have high skills and qualifications; in some sense, working in a high-productivity occupation is a characteristic of the individual. In our model, our identifying assumption implies that equally productive workers could work in different occupations for reasons uncorrelated with their wages. Thus, in our model, it is a feature of the job that they match to in the labour market. It should be noted that of course any individual-specific and time-invariant component of one's occupation quality will remain a part of the individual fixed effect.

Why do we believe that our interpretation of occupations is preferable to the AKM interpretation? The main reason is simply that people do change their occupations when they move between firms, and this is something that is not accommodated neatly in the AKM setting. While there are many people who remain in the same occupation throughout their working life, there are many who are mobile between different occupations. Work by Acemoglu and Autor [2011] has also pointed out that different occupations requiring performance of different kinds of tasks may offer different levels of compensation. Thus, the standard AKM interpretation may understate the degree of wage heterogeneity experienced by observationally equivalent workers.

Job fixed effects capture the time-invariant component of log-pay due to job specific heterogeneity conditional on worker quality within said job. If otherwise similar workers are in job A and B, but those in job A receive higher pay, this will be reflected in the job fixed effect. As discussed, relative to firm fixed effects, job fixed effects allow firms to pay different wage premia to those working in different occupations.

2 Data and empirical setting

To apply the econometric model described above, we require panel data on the wages, firm, occupation and experience of the universe of workers. We use data from France, one of the biggest countries with such administrative data available[7]. To test the validity of our results outside France we replicate the main results using analogous matched employee-employer administrative data from Germany. German data has the disadvantage that wage information is top-coded, and thus is not informative about wage variance in particularly high-paying occupations. For this reason, we do not focus on the German results in the main body of the paper but rather use them to qualitatively give some evidence of the external validity of our main results.

2.1 French administrative data

We start with the French administrative datasets "Base tous salariés" (hereafter BTS) Insee [2024]. These are annual cross-sectional matched employer-employee datasets covering the universe of French workers. A unique feature of these datasets is that they provide information not only on the worker in the current year t, but also on the worker in the previous year $t-1$ as well. Babet et al. [2022] propose a methodology to chain these repeated cross-sections to create a quasi-panel tracking individuals over time by matching data on individuals in time t in the year t dataset and in time t in the year $t+1$ dataset. Details on the methodology used to create the panel are described in Babet et al. [2022][8]. This methodology allows us to match over 95% of individuals each year. One major limitation of this approach is that those who are out of the labor force (or in government employment) for more than one calendar year cannot be matched over time and will instead be given a new individual identifier.

The pseudo-panel constructed from the BTS data provides information on individual characteristics (age, gender), occupation, granular location at the commune level[9], detailed pay and hours worked, and firm. Relative to other administrative datasets this is a particu-

[7] US matched employee-employer data does not include information on occupations.

[8] We are indebted to the public good provided by Babet et al. [2022] who have kindly made code available to allow us and other researchers to follow their methodology easily.

[9] There are 34,955 communes in France.

larly rich set of information, that doesn't suffer from top coding of salaries.

We follow the literature in restricting the sample to only private sector workers, working at least 13h a week on average, working in metropolitan France, earning more than 2/3 the annualised minimum wage and 80% of the hourly minimum wage, and finally earning less than 1,000,000 euros a year in 2020 prices. Finally, we restrict the analysis to men[10] aged between 25 and 60 inclusive and focus on the period from 2015 to 2019, stopping before the onset of the COVID-19 pandemic. We also provide a decomposition for the earlier period 2010-14 and a decomposition considering only women in appendix section A.2.1 and present summary statistics for the period in the summary statistics table below.

In our main analysis we focus on the four-digit "PCS" (professions et catégories socio-professionnelles) classification which gives 430 unique occupations. Examples of occupations are: midwives, professors and lecturers, surveyors and topographers, and qualified roofers. A worker's occupation information is collected from compulsory monthly employer surveys. This monthly information is then aggregated to the spell level and we follow the literature and Babet et al. [2022] in taking the highest earning job in a given year. Reported occupation codes are cross-checked against the reported occupation description using specialist software by INSEE. In 90% of cases the codes agree and various correction processes are implemented to recover the remaining cases.[11] A possible threat to our results is measurement error in occupation classification. Given that only the (wage-weighted) most frequent monthly occupation is taken, we expect this measurement error to be minimal; however, we additionally address this concern in three ways. First, we find qualitatively similar results when we only use variation in occupations due to firm-changes which are less likely to be due to measurement error. Second, we find qualitatively similar results using German administrative data which uses an entirely different data collection process. Third, measurement error is less likely to be present at coarser definitions of occupation (3, 2, or 1 digit) and although results differ due to coarseness (as one might expect) qualitatively the main take-aways remain.

Our main measure of earnings is annual earnings; we replicate our main decomposition

[10] We focus on men for computational reasons. We consider women separately in appendix A.2.1, and find no qualitative or quantitative differences.

[11] More information can be found in an INSEE "Statistical Mail" found at https://www.insee.fr/fr/information/3647029?sommaire=3647035.

using hourly earnings in appendix A.2.1 and find qualitatively similar results. In our main results, we use residualised log annual wages, constructed by first regressing log annual earnings against a cubic wage profile and year fixed effects and taking the residual. We use firm identifiers (SIRENE firm numbers) supplied in the BTS data. We define a job as a unique occupation-firm pair; thus, our job fixed effects can also be interpreted as firm-occupation interaction effects.

Table 1 shows some summary statistics from the resulting main sample using data from 2015-19.[12] After implementing the restrictions our main sample covers 48 million observations consisting of 14.2m workers, 1.3m firms, and 4.5m jobs. Once we consider the leave-one-out (LOO) connected set for the worker-jobs model, we have significantly fewer observations (around 66% of the overall total). The LOO connected set for the jobs model which underlies our main specification contains 57% of workers, 21% of firms, and 18% of jobs in the full data. This compares to 81% of all observations, 69% of workers, 30% of firms, and 58% of jobs in the full data for the LOO of the worker-firms model that would underlie a standard AKM decomposition.

The average wage for the connected and LOO sets (for both the jobs and firms models) are marginally higher (0.04 log points for annual wages, 0.02 log points for hourly wages), while the variance is lower (by approx 0.02, or about 8%).

We also provide summary statistics on the number of moves we observe in the data. In column 1 of table 1, we report the total number of times workers are observed to change either their firm or occupation, as well as the number of moves that are between firms, the number of moves between occupations and the number of moves between both firms and occupations. In the full data, we find that 14.5% of observations involve moves. Of these moves, 51.6% involve moves between firms, 79.0% involve moves between occupations and 30.5% involve changes in both firms and occupations. In the LOO, 14.0% of all observations involve moves, of which 25.4% are moves within occupations between firms, 42.9% are moves within firms between occupations, and 31.8% are moves between both firms and occupations.

[12]Following the literature, to address concerns about potential changes in job fixed effects over time, we focus on a period of 5 years. To assuage concerns that our results are driven by the specific period we study, we also report results on an earlier period, 2010-14, to show that our results are similar in that earlier time period. Summary statistics for this period is reported in table 4 in appendix A.1.

Table 1: Summary statistics

	Full data	Firms connected set	Firms LOO set	Jobs connected set	Jobs LOO set
2015-19					
N obs	48,012,526	40,777,782	38,749,595	35,928,241	31,529,737
N workers	14,247,535	10,282,172	9,781,898	9,059,072	8,086,939
N firms	1,251,983	645,578	376,323	571,440	266,821
N jobs	4,479,788	3,166,687	2,603,905	2,040,895	819,020
Mean log annual wage	10.39	10.42	10.43	10.42	10.43
Var log annual wage	0.25	0.24	0.23	0.23	0.23
Mean log hourly wage	2.96	2.97	2.98	2.97	2.98
Var log hourly wage	0.21	0.2	0.2	0.2	0.19
Var residualised log hourly wage	0.24	0.22	0.22	0.22	0.21
N moves	6,966,208	3,406,086	3,129,298	5,689,874	4,427,720
N firm moves	3,591,255	3,406,086	3,129,298	3,251,625	2,530,089
N occ moves	5,503,513	2,673,604	2,425,116	4,379,023	3,303,582
N firm + occ moves	2,128,560	2,032,493	1,836,961	1,940,774	1,405,951

Notes: This table shows the summary statistics from the main sample used in this paper. The underlying data is from yearly BTS data files, French administrative matched employee-employer data. Individuals are mapped over time using the procedure and kindly provided programs in Babet et al. [2022]. Sample construction and restrictions are discussed in the text.

3 Worker-job log wage decomposition

Decomposing log wage variance using equation 2 we find that: Individual heterogeneity contributes only 49% of total wage variance, the variance of job fixed effects contributes 11%, while the covariance between the two fixed effects accounts for 21%. These results are reported in the final column of table 2 which corrects for limited-mobility bias using the leave-one-out correction [Kline et al., 2020]. This contrasts to the 57%, 6% and 9% respectively according to the corrected AKM model (column 4, table 2); the individual component has fallen by 8pp while the share due to job heterogeneity has increased by 4.2 pp and the share due to sorting has more than doubled. We also find that the worker-job model explains more of the variation overall with variance attributed to the residual falling by 8 percentage points.

Table 2: Baseline decomposition results

	Firm AKM		Firm KSS		Job AKM		Job KSS	
	Var	Prop	Var	Prop	Var	Prop	Var	Prop
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worker ($Var(\alpha)$)	0.159	0.779	0.118	0.574	0.127	0.639	0.098	0.493
Firm ($Var(\psi)$)	0.021	0.101	0.013	0.064				
Job ($Var(\lambda)$)					0.033	0.165	0.021	0.106
Sorting ($2 \times Cov(\alpha, \lambda)$)	0.004	0.022	0.019	0.091	0.021	0.104	0.042	0.21
Error ($Var(\varepsilon)$)	0.02	0.099	0.056	0.271	0.018	0.093	0.038	0.192
Residualised $Var(Y)$	0.205	1	0.205	1	0.198	1	0.198	1

Notes: This table presents decompositions of the residualised log wage for the worker-firm fixed effects model in columns 1-4, and the worker-job fixed effects model in columns 5-8. All decompositions are estimated on French matched employee-employer administrative data from 2015 to 2019, using the largest connected leave-one-out set. Columns 2-3 and 6-7 show the plug-in decompositions for the worker-firm and worker-job models respectively as variances (columns 1 and 5) and shares of total residualised log wages (columns 2 and 6). Columns 3-4 and 7-8 show the variances corrected due to Kline et al. [2020] for a worker-firm model and a worker-job model respectively. Column 3 and 7 shows the variances and columns 4 and 8 present the shares.

We find that naive decompositions overestimate the proportion of log-wage variance due to individual and firm (job) heterogeneity at the expense of that due to sorting — as is typically found in the literature.[13] However, we find similar patterns contrasting the uncorrected decompositions of the firms and jobs models (columns 2 and 6) — the role of individual heterogeneity is diminished while the share of job heterogeneity and sorting increases substantially.

The reduction in the share of variance due to individual fixed effects is about 14% or 8.1 pp. The individual fixed effects account for time-invariant individual differences in their resulting pay. These differences could include differences in cognitive ability, inherent worker skills or education, long-term health conditions and possible discrimination due to immutable personal characteristics. Our results suggest that some of this was due to the correlation between the individual fixed effect and the occupation that a worker is in. Intuitively, worker-firm relative to the more flexible worker-job specification will overestimate variance due to worker heterogeneity if individuals who tend to work in high-wage firms also tend to work in high-wage occupations. This is because when the job fixed effect is not included in the regression in column (4), it shows up as part of the omitted variable bias in the estimate of α_i in the worker-firm model if there is residual covariance between α_i and $\lambda_{J(i,t)}$ after accounting for the firm workers work at. This bias component is counted as part of the

[13] We provide an overview of select results from the literature in appendix A.7.

individual fixed effect in the variance decomposition of the misspecified model and must be larger than the actual variance of α_i if the covariance between the individual fixed effect and the residualised, weighted sum of job fixed effects across an individual's career is positive.[14] This argument is formalized in appendix A.6.

We find a significantly larger role for job-level heterogeneity and worker-job sorting (accounting for 31.6% of total log wage variance) than worker-firm decompositions attribute to firm-level differences and worker-firm sorting (accounting for 15.5% of total log wage variance). This presents indicative evidence that occupation-level heterogeneity and sorting may be playing an important role and that firms pay varying wage premia to different occupations/ tasks. To provide more direct evidence on these dimensions we first decompose the variance due to job heterogeneity and worker-job sorting into that which can be attributed to within-occupation cross-firm differences and that due to differences across occupations using the law of total variance. This exercise is described in subsection 3.2. Second, by comparing the within-firm variance in job fixed effects across firms and the difference between additive and multiplicative occupation-firm models, we provide direct evidence on whether wage premia varies within a firm between occupations. These results are reported in subsection 3.3.

3.1 Specification tests and robustness

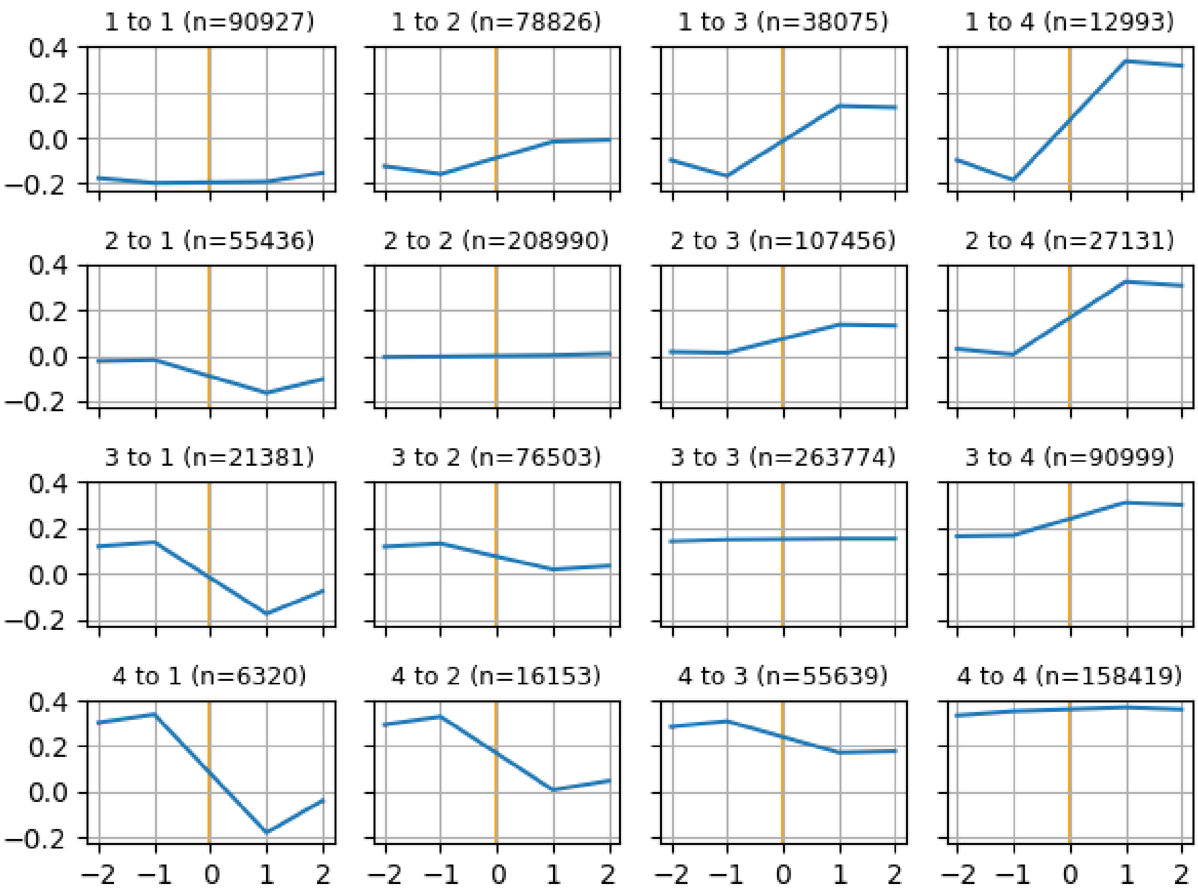
3.1.1 Testing the identification assumption

In this section, we implement the event-study "tests" of the identification assumption discussed and explained in section 1.2 following Card et al. [2013]. We categorise jobs individuals move into or out of into four earnings quartiles using the leave-out job-specific mean fixed effect. Figure 1 then plots how average wages move around the event of a job switch between each of the four categories. Focusing on the most extreme moves from the first to the fourth quartile and from the fourth to the first, there is a clear symmetry in the impact. The other cells, although less extreme, show a similar symmetry. We produce an analogous diagram (figure 9 in appendix A.4) categorising jobs using average job wages in appendix A.4. Sym-

[14] This result is analogous to the biases due to what they term 'hierarchy effects' in the analysis of industry pay premiums in Card et al. [2024].

metric wage changes when moving between different quality jobs, stable wages when moving between similar quality jobs, and stability in the years around a move all give credence to the key identifying assumptions and follow a pattern of similar results in the worker-firm literature Bonhomme et al. [2023], Card et al. [2013, 2018].

Figure 1: Event study around job moves, clustering by leave-out job mean wages



Notes: This figure shows the impact on average wages around the event of job movement. Each cell shows the average wage change associated with a movement event from a quartile to another quartile of the average job fixed effect distribution. Following Card et al. [2013], we cluster jobs into quartiles by computing the mean leave-out job fixed effect within the job excluding. Only those who remain in their old job for two years before and their new jobs for two years after the move event are included. The number of switchers in each cell is given in the cell title.

An advantage of considering job as opposed to firm heterogeneity is that we can leverage variation due to movers across occupations as well as firms. Focusing on the largest leave-one-out connected set in our main sample period, we find that 43% of moves are across occupations within a firm.

One concern with using job moves instead of firm moves might be that occupational changes within the firm could not be as conditionally exogenous as moves across firms, which had been extensively studied in other AKM studies. Occupational changes within firms might represent promotion or demotion events related to match effects, e.g. due to better firm learning about the worker's abilities, or human capital accumulation beyond what is captured by the inclusion of a cubic age profile. We investigate this potential issue in two ways. Firstly, and most directly, we can estimate our baseline decomposition results on a sample that doesn't include those who move within a firm — and therefore abstracts away from such variation. We perform this exercise in the appendix sub-section A.2.4 and show that the main results remain unchanged.

Second, we perform Card et al. [2013] style event studies around moves between occupations within firms, excluding moves between firms more commonly considered in other AKM studies. We present the result in figure 10 in appendix A.4. We observe symmetric wages changes in moves between estimated job effect clusters even when only considering moves across occupations within firms, leading us to think that promotions are not particularly problematic violations of the identifying assumption.

3.1.2 Robustness

In the appendix section A.2 we show that the qualitative conclusions reached above are robust to considering (i) hourly wages as opposed to annual earnings, (ii) using the entire connected set and a homoskedastic decomposition correction, (iii) varying granularity of occupations (1,2, and 3 digit), (iv) an earlier time period (2010-2014), (v) only considering cross-firm moves, (vi) women workers.

We also investigate whether our results are robust to excluding moves within firms across occupations, which might be problematic if they were less likely to be exogenous than moves across firms which have been widely studied in previous papers in the AKM literature. We compare the baseline decomposition to an alternative decomposition dropping all workers who move within firms in table 8. We find that excluding such moves reduces the share of wage variance attributed to worker heterogeneity and increases the share due to job variance. The share attributed to sorting between worker and jobs also increases by 4.2 pp.

To analyse whether our results are specific to the French setting, we perform an analogous decomposition using data from German social security records. These results are presented in full in appendix A.3. In general, we find that the German data is quite comparable to the French data in terms of the size of the population, and the number of firms. We also find a qualitatively similar decomposition with worker heterogeneity accounting 42% of total wage variance, jobs 18% and sorting 21%. Contrasting this with the French results, the role of individual heterogeneity is lower by 7pp, and the role of job heterogeneity is higher by 7pp. The role of sorting is similar in both contexts.

3.2 Between and within-occupation decompositions of the variance of job fixed effects and worker-job covariance

We find that job-level heterogeneity and worker-job sorting are both quantitatively more important than firm-level heterogeneity or worker-firm sorting in residualised log-wage decompositions. As jobs are defined by firm-occupation pairs, a key question is whether heterogeneity in the job fixed effect is due primarily to differences between firms or between occupations. Answering this question allows us more precisely to consider whether firm-level or occupation/task-level differences are of first-order importance. To get at this, we further decompose the variation of job fixed effects, and the covariance between job and worker fixed effects, into between- and within-occupation components, using the law of total variance. Let $\mathbb{E}_o(\cdot)$, $\mathbb{V}_o(\cdot)$, and $Cov_o(\alpha_i, \cdot)$ denote the mean, variance and covariance with α_i of the random variable within the partition indexed by occupation denoted by o. Using the law of total variance equation 3 presents the decomposition of job variance into between and within occupation components, while equation 4 does the same for the decomposition of covariance.

$$\mathbb{V}[\lambda_j] = \underbrace{\mathbb{V}[\mathbb{E}_o[\lambda_j]]}_{\text{Between } o} + \underbrace{\mathbb{E}[\mathbb{V}_o[\lambda_j]]}_{\text{Within } o} \tag{3}$$

$$Cov[\alpha_i, \lambda_j] = \underbrace{Cov[\mathbb{E}_o[\alpha_i], \mathbb{E}_o[\lambda_j]]}_{\text{Between } o} + \underbrace{\mathbb{E}[Cov_o[\alpha_i, \lambda_j]]}_{\text{Within } o} \tag{4}$$

The computation of the terms of the decomposition is complicated by the fact that the fixed effects are not observed, and only estimated with error. As such, naive plug-in estimators of the terms of the decomposition will be biased. First, $\mathbb{V}_o[\hat{\lambda_j}]$ is a biased estimator of $\mathbb{V}_o[\lambda_j]$, and thus, taking means over the occupations of the naive plug-in conditional variance estimator will also lead to a biased estimate of the latter term of equation 3. Second, while $\mathbb{E}_o[\hat{\lambda}_j]$ is an unbiased estimator of $\mathbb{E}_o[\lambda_j]$, it is still estimated with error. Therefore, plugging in the former for the latter will lead to a biased estimator of the former term for reasons similar to those which lead to limited mobility bias. However, if occupations are sufficiently large, we should expect the bias to be small since $\mathbb{E}_o[\hat{\lambda}_j]$ should estimate $\mathbb{E}_o[\lambda_j]$ with relatively low error.[15]. Similar concerns plague the estimation of the terms of equation 4.

We report two estimates of the terms of equations 3 and 4. First, we compute the between component of the decomposition using a naive plug-in estimator, and report the approximate within component as the differences between the KSS corrected estimate of aggregate variance and the approximate plug-in estimate of the between component. Second, we compute KSS corrected estimates of $\mathbb{V}_o[\lambda_j]$ and $Cov_o[\alpha_i, \lambda_j]$ for each occupation.[16] Since the KSS corrected estimates are unbiased for the true within-occupation variance of job fixed effects and covariance between individual and job fixed effects, we estimate the within component as the across-occupation employment-weighted mean of the KSS corrected conditional variances and covariances. The between component is computed as the difference between the KSS corrected aggregate means and the estimated within component.

[15]This naive plug-in estimator is analogous to the "bottom-up" approach pursued in Card et al. [2024] to estimate industry wage premiums. As a result, analogous conditions need to be satisfied in Card et al. [2024], that is occupations need to be sufficiently large. Note that this implies that a "bottom-up" approach to estimating corrected firm fixed effects is unlikely to work as although there are on average over 90,000 observations per occupation there are only around 30 observations per firm.

[16]To do this, we note that all conditional variances and covariances of fixed effects can be represented in matrix form as:

$$\theta = \beta' A \beta$$

Our approach is to find the positive semi-definite matrix A, which corresponds to the conditional variance or covariance. We describe this in more detail in appendix A.8.

Table 3: Between-within occupation decompositions of the variance of job fixed effects and worker-job covariance

	Between (Share)	Within (Share)	Total
Approx between			
Job FE ($Var(\lambda)$)	0.010 (0.48)	0.011 (0.52)	0.021
Indiv, Job Cov ($Cov(\alpha, \lambda)$)	0.017 (0.81)	0.004 (0.19)	0.021
KSS corr. within			
Job FE ($Var(\lambda)$)	0.010 (0.48)	0.011 (0.52)	0.021
Indiv, Job Cov ($Cov(\alpha, \lambda)$)	0.017 (0.81)	0.004 (0.19)	0.021

Notes: This table presents the decomposition of the variance of job fixed effects and the covariance between worker- and job- fixed effects (from columns 8 and 9 of table 2) into between- and within-occupation components, computed either with a naive approximation of the between-occupation variance (panel 1) and a procedure to calculate within-occupation variance by averaging over KSS-corrected conditional variances (panel 2).

Table 3 presents the result of these calculations. We find that the two methods coincide (to three decimal places), suggesting that the noise in the within-occupation means does not lead to substantial bias in the estimation of the between term of the decomposition. We find that just under half of the total variance of job fixed effects can be explained by between-occupation variation, with the other half being explained by within-occupation, between-firm variation. The within-occupation between-firm component amounts to 0.011 or about 5.2% of total residualised log wage variance, while the between-occupation component amounts to 0.010, or about 4.8% of total residualised log wage variance.

It is interesting to compare these magnitudes to the findings from the AKM model; there, the variance of firm fixed effects in the AKM decomposition in table 2 is about 0.014 or 6.4% of total residualised log wage variance. This number is similar to the 0.011 figure for within-occupation between-firm variation, suggesting that the increase in the variance of job fixed effects is driven primarily by between-occupation variation. The variance of firm fixed effects in the AKM model is slightly larger than the share due to within-occupation variation in the job model. This is possibly because worker-firm models will conflate firm FE with occupation fixed effects if there is some correlation between the wage premia at the firm and the wage premia at the occupation.

Similarly sorting between workers and jobs can be decomposed into within and between occupation components. The between occupation component is given by the covariance between mean worker type within an occupation and mean job fixed effect within that occupation. Whereas the within occupation component is given by the expected covariance between workers and jobs in a given occupation, across all occupations. We find that 81% of the variation due to sorting can be explained by sorting between occupations, which reflects the sorting of higher-paid workers into higher-paying occupations, while only 19% is attributed to sorting within occupations between firms. This result suggests that a large part of the sorting documented in the AKM model might be due to the way that occupations cluster between firms.

We perform a similar decomposition using the German social security records; this robustness exercise is reported in appendix A.3.4. We find that in the German setting, 70% of the variance in job fixed effects is due to between-occupation variation and almost all of the covariance between individual and job fixed effects is due to sorting between occupations, and almost no variance can be attributed to sorting within occupations between firms.

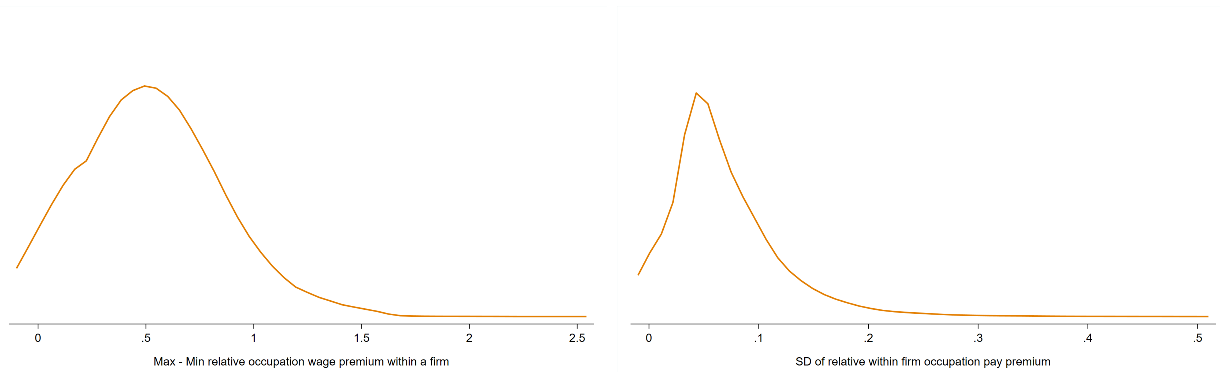
Our results thus offer a different perspective on log wage inequality than most of the literature. Results from standard AKM regressions have tended to primarily emphasise the role of individual heterogeneity and secondarily the role of firms. In contrast, our results show that the occupational structure plays a large role in the story of wage inequality. Our results suggest that sorting across occupations is quantitatively more important than sorting across firms in explaining wage variance. This finding may cohere with the emphasis on occupations typically found in analyses of wage inequality focusing on tasks (e.g. Acemoglu and Autor [2011]).

3.3 Direct evidence on varying wage premia within firms

Using the estimated worker and job fixed effects, we analyse the extent to which firms do indeed offer different wage premia to workers in different occupations. To do this, we study the dispersion in the distribution of job fixed effects $\widehat{\lambda}_j$ within a firm f having removed aggregate occupation variation. We plot the cross-firm distribution of $\widehat{\lambda}_j - \widehat{O}_o$ where $\widehat{O}_o$ are estimated occupation fixed effects from a regression of estimated job fixed effect on occupation dummies. Figure 2 plots two statistics from this distribution; panel (a) plots the cross-occupation within-firm range and panel (b) plots the cross-occupation within-firm standard deviation. Both statistics are measured in terms of log wages and so a range of one implies that having removed firm- and occupation-specific means, this firm pays an occupation twice as much as another. Similarly, a standard deviation of 0.1 in panel (b) can be interpreted as within a firm across occupations that occupations paying one standard deviation more pay on average 10% more. Both measures show considerable variation that would otherwise be unexplained in worker-firm decompositions.

Figure 2: Distribution of the variance of job fixed effects across occupations within firms

(a) Within-firm cross-occupation range | (b) Within-firm cross-occupation variance



Notes: This figure shows the within-firm, cross-occupation dispersion in wage premia removing aggregate cross-occupation effects. Both figures show density plots over firms, in the left-hand panel the cross-occupation range (max-min) is displayed and in the right-hand panel, the cross-occupation variance is displayed. The sample is restricted to firms with 10 or more employees. To calculate the statistics graphed, estimated job fixed effects are residualised on firm and occupation effects.

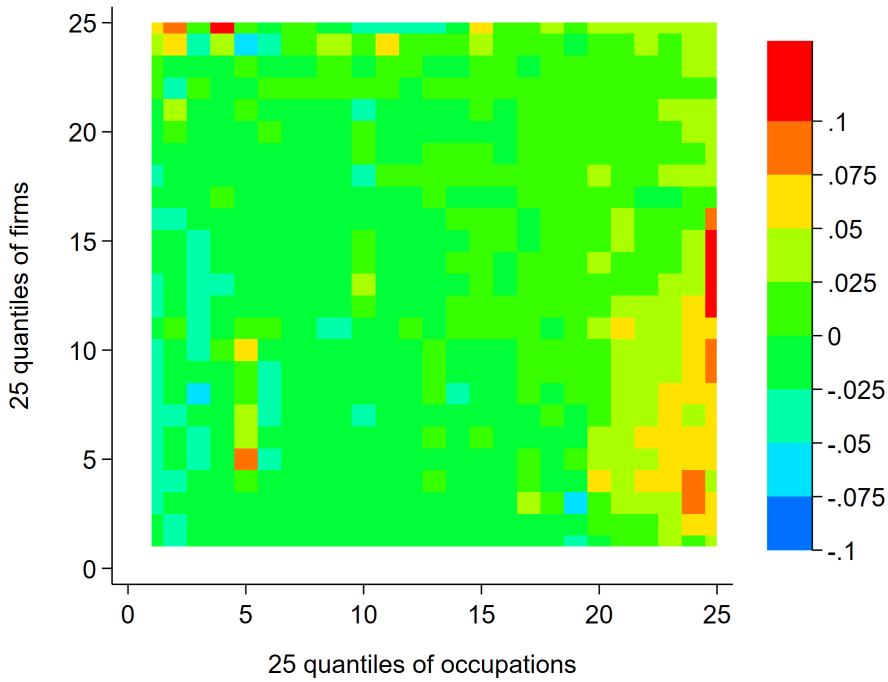
Another possibility is that while firms may offer an occupational premium to workers in different occupations, they also impose a constant firm premium on all workers within a firm. If so, we can approximate the job fixed effect estimated in our model by an additive specification given by equation 5, where $occ(i,t)$ denotes the occupation worker i is in at time t and $F(i,t)$ denotes the firm the worker is in at time t (studied in Torres et al. [2018]).

$$\lambda_{J(i,t)} = \phi_{occ(i,t)} + \chi_{F(i,t)} \tag{5}$$

To test this interpretation, we estimate regressions of the form $\ln(w_{it}) = \alpha_i + \phi_{O(i,t)} + \chi_{F(i,t)} + \varepsilon_{it}$, where ϕ_O are occupation fixed effects and χ_F are firm fixed effects. Then we can study the difference between these estimated fixed effects and our previously estimated job fixed effects $\lambda_{J(i,t)}$. We interpret this difference as an occupation-job specific match effect beyond what is predicted by log linearly additive firm and occupation fixed effects. If firms do not pay differential wage premia across occupations, $\Delta = \lambda_J - (\tau_O + \xi_F)$ will be small and not systematically related to any firm or occupation characteristics. A positive match effect of e.g. 0.1 implies that the premium offered by the particular job is roughly 10% greater than the premium implied by an additive occupation and firm fixed effect (and vice versa).

In figure 3, we produce a heat map of the difference between the job fixed effect in our main model and the sum of the occupation and firm fixed effects in the auxiliary model, that is of match effects. We plot this for 25 quantiles of occupations ranked by the average log hourly pay in the occupation and 25 quantiles of firms ranked by the average log hourly pay in the firm. Thus, a positive number (e.g. 0.05) can be interpreted as follows: if we naively replace the job fixed effect by the sum of the two fixed effects in our predicted values in our main model, we would underestimate the log wage by approximately 5%. Deep and light green tiles represent quantile combinations where the error is less than 0.05 log points in either direction.

Figure 3: Contribution of worker-job match effects by firm and occupation mean pay quantiles



Notes: The figure shows a heatmap of the difference between the estimated job fixed effect in our main specification, and the sum of a firm fixed effect and an occupation fixed effect in an auxiliary specification, which we interpret as match effects between occupation and firms. Green squares imply that if the job fixed effect were replaced by the additive firm and occ fixed effects, the resulting predicted log wage would be within 0.05 log points of the true predicted log wage with job FE. Red implies that using the additive specification would underpredict the true wage by over 0.1 log points and blue implies that overprediction by over 0.075 log points. The x-axis indexes 25 quantiles of occupations, sorted by the mean wage within the occupation, and the y-axis indexes 25 quantiles of firms, sorted by the mean wage within the firm.

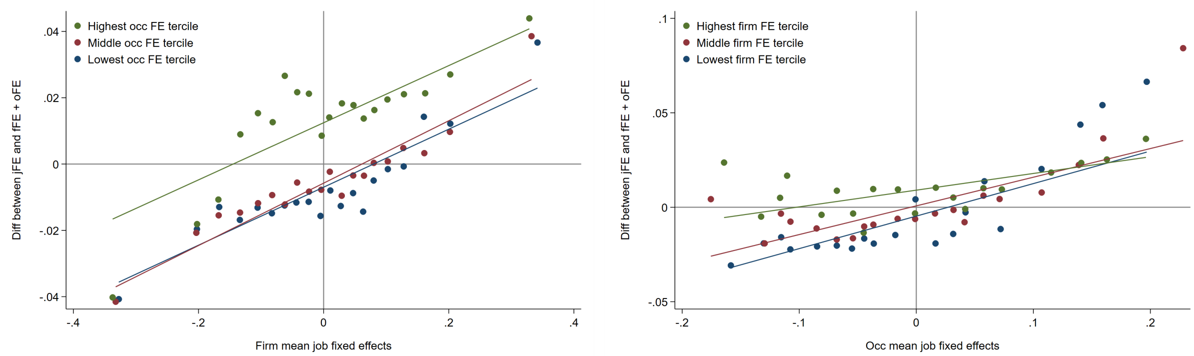
This figure shows that the linear form performs quite well in general, but seems to overestimate the wages of workers in low type occupations at some firms (as these have a negative match effect), and severely underestimate the wage of high type occupations especially at relatively lower wage firms. The problem seems especially acute for high wage occupations at low wage firms and low wage occupations at high wage firms, leading to mistaken predictions by over 10%. This leads us to conclude that the log additive firm and occupation fixed effect specification fails to fully capture the interaction between occupation and firms in wage determination.

To analyse this dynamic further, we produce two complementary figures. In each figure, we plot the estimated firm-occupation match effect, that is, the difference between the unrestricted job FE and the additive firm and occupation fixed effects, against the mean job fixed effects within the firm and the occupation respectively, separating the data into terciles ordered by the within occupation and firm means respectively.

Figure 4: Estimated firm-occupation match effects by firm and occupation fixed effect terciles

(a) Match effect by firm mean JFE for three occupation JFE terciles

(b) Match effect by occupation JFE for three firm JFE terciles



Notes: This figure plots the distribution of the difference between the estimated job fixed effect in our main model, and the sum of estimated firm and occupation fixed effects in an auxiliary regression. Panel 4a plots the differences by firms, sorted by the mean JFE within the firm for three terciles of occupations, while panel 4b plots the differences by occupation, sorted by mean occupation JFE, for three terciles of firms.

Figure 4 shows binned scatterplots of match effects which emphasise the importance of match effects for job-level pay premia[17]. Panel 4a shows that match effects increase from negative to positive as mean pay premia within firms increase. This relationship, however, is mediated by occupation terciles. Low- and medium-wage occupations have a negative match effect at most firms, regardless of the firm-level pay premium. On the other hand, high-wage occupations have a positive match effects at most firms, even those who offer on average below-average wage premia. Panel 4b shows the analogous results when plotting match effects over occupation average job fixed effects and cutting by firm occupation fixed effect terciles. We find that lower occupation fixed effects occupations consistently have smaller negative match effects than higher wage premium occupations.

Our results support papers that argue that firms are likely to offer different wage premia to different workers, such as Haanwinckel [2023], and Bloesch et al. [2022]. Haanwinckel [2023] argues that firms should pay larger premia to workers who perform tasks which the firm use more intensively, and in a similar vein, Bloesch et al. [2022] show that workers which

[17]The scatterplots can be interpreted as similar to the data in figure 3, 'sliced' into a 2-D graph by reducing the dimensions of one of the axes.

have more hold-up power within a firm should be paid more. Panel 4a shows that these jobs are likely to be more highly paid and skilled occupations. Assuming a proportional firm wage premium leads to the systematic overestimation of the pay of low-wage occupations and the underestimation of the pay of high-wage occupations at high productivity firms.

4 Heterogeneity across labor markets

We have thus far only considered average decompositions across potentially many heterogeneous labor markets. However, it may be the case that this average does not accurately represent all labor markets. Additionally, by considering heterogeneity across labor markets we can begin to understand what characteristics are associated with a greater role for firms or occupations, or sorting or individual heterogeneity. In this section, we consider heterogeneity of the sources of inequality on two dimensions: (i) within occupations, and (ii) within commuting zones.

Within occupations, we decompose the variance of log wages into the variance of the individual FEs, the variance of the job FEs (which within occupations reduce to firm FEs) and the covariance between the two components. As with the aggregate decomposition and as discussed in Andrews et al. [2008] and Kline et al. [2020], naive plug-in estimates of the conditional variances of the fixed effects are biased. Instead, we compute KSS-corrected estimates of the conditional variances, which are unbiased[18]. As with the aggregate decompositions, we are not able to compute the KSS corrections exactly due to computational limitations. Thus, the estimates presented are approximations. Unlike with the aggregate decomposition however, an additional challenge is that relatively smaller sample sizes may result in a less precise approximation; the degree of approximation error which may increase as the conditional group gets smaller. This approximation may lead to attenuation in the relationships that we document. Therefore, such relationships should be considered a lower-bound.

Within commuting zones, we can further decompose the within-commuting zone variance of job fixed effects into between and within occupation components. We estimate the between-occupation component by taking the variance of within-occupation fixed effect

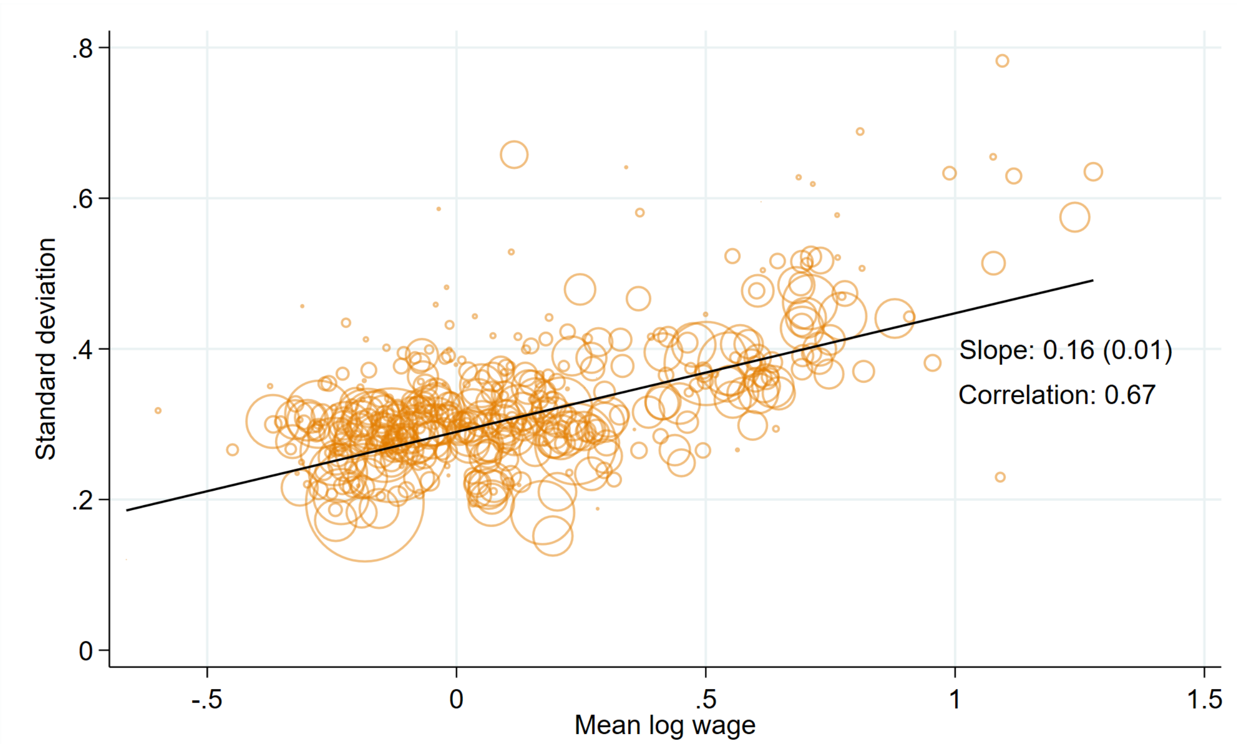
[18] See appendix A.8 for further details.

means, which will lead to a degree of incidental parameter bias if commuting zone-occupation cells are particularly small.

4.1 Heterogeneity across occupations

The quantity of wage dispersion within occupations varies considerably across occupations. Figure 5 displays this and depicts the stylised fact that in the French labour market, the variance of log hourly wages is co-increasing in the mean log hourly wage within an occupation. Figure 5 shows a clear log-linear relationship between the mean and variance of residualised log earnings across occupations.

Figure 5: Within occupation mean vs SD of residualised wages



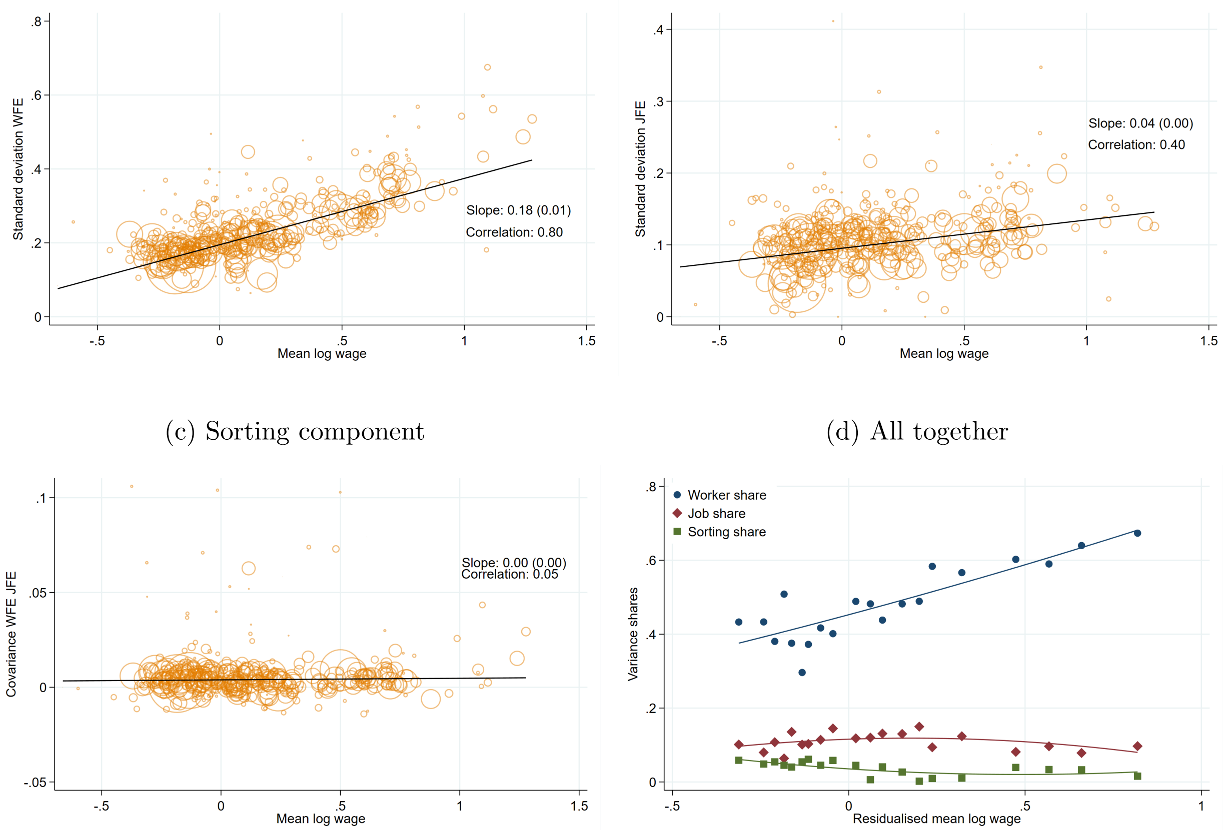
Notes: This figure plots the mean log-wage (x-axis) and the standard deviation of residualised log-wages with respect to a cubic age profile and year fixed effects (y-axis) within each occupation for all occupations in the data. Each bubble represents a particular occupation and the larger the bubble is, the more workers there are in the occupation. A linear trend line is drawn, with slope 0.16.

What accounts for the greater wage variance in high-wage occupations relative to lower wage occupations? To understand this, we perform a variance decomposition of residualised log wages in each occupation into an individual component, a job component (which within occupations reduces to a firm fixed effect) and a sorting component. We plot the log-variance due to each component for each occupation in figure 6, panels (a) to (c). In panel 6d, we plot the variance of each component as a share of total log-wage variance. We find that while the variances of worker and job fixed effects are increasing in the within-occupation mean wage, there is no relation between the sorting component and mean occupation wage. Analysing the variance components as shares of total residualised wages, we find that worker heterogeneity is the most important determinant of total log wage variance, while the variance of job fixed effects (within occupations) and sorting are of relatively similar magnitude.

Despite worker and firm heterogeneity being relatively larger in high wage occupations, we find that the covariance between worker and firm within occupations fixed effects is not also increasing in mean occupation wage. Higher paid workers do not seem to be more likely to sort to higher-paying firms within higher-paying occupations.

Figure 6: Log wage variance decomposition within occupations (shares)

(a) Worker component (b) Job component



(c) Sorting component (d) All together

Notes: Panel 6a plots the standard deviation of worker fixed effects in each of the 430 occupations. Panel 6b plots the standard deviation of within-occupation job fixed effects within each occupation; since a job is identified as an occupation-firm pair, the variation is from firms offering different wages for the same occupation. Panel 6c plots the correlation between worker and firm fixed effects for each occupation. In panel 6d, we plot the variance of each component as a share of total log-wage variance. The occupations are ordered by the mean log-wage within the occupation on the x-axis.

These heterogeneity results show that in higher-wage occupations, firm heterogeneity matters more although sorting of workers into firms does not. Additionally, individual heterogeneity is increasingly important and the unexplained/ control components of the variance decomposition are increasingly unimportant. These results imply that models emphasising firm differences (derived from varying firm-level productivity or monopsonistic power) are somewhat more applicable to low-wage occupations, whereas individual ability is more heterogeneous (or is at least more wage-elastic) in high-wage occupations.

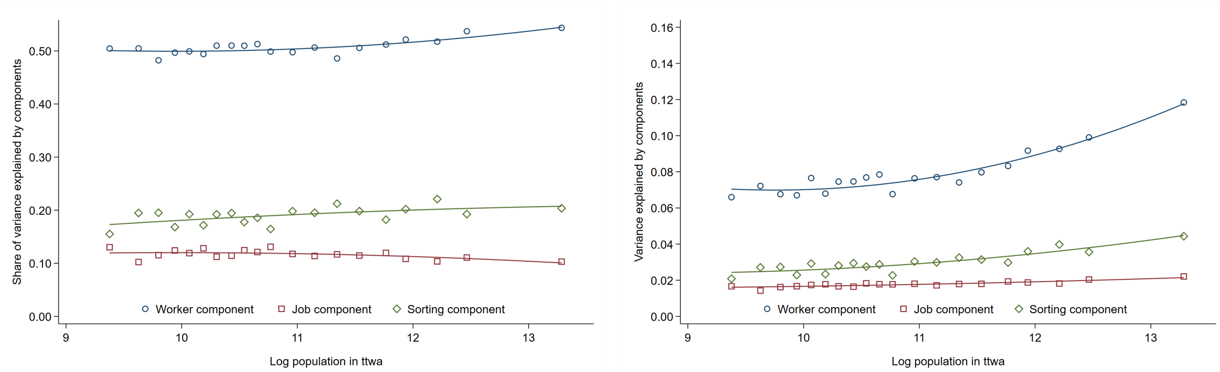
4.2 Heterogeneity by local labour market

Secondly, we turn to characterize labour markets by geography. We split (metropolitan) France into its 340 commuting zones as defined by the French statistical institute INSEE.[19] We first apply the main log-wage decomposition to each commuting zone, splitting log wage variance into a component due to the variance of individual fixed effects, a component due to the variance of job fixed effects, and a component due to the covariance between the two. We relate each of these components to (log) commuting zone population to characterise how the importance of each element relates to the population of the commuting zone. Unlike previous literature [Dauth et al., 2022], which estimates separate fixed effects regressions for each commuting zone, we instead compute corrections for the naive plug-in conditional variance estimates following the methods in Kline et al. [2020].

The results of this exercise are reported in the graphs in figure 7. We find that both the variance of individual fixed effects and sorting are increasing in local labour market size, but the share due to job fixed effects is stagnant[20]. This result coheres with a previous result in Dauth et al. [2022] which finds that sorting to firms accounts for more variance in larger labour markets in Germany. This finding could be explained by the value of search being greater in larger cities due to scale effects, as was found in Petrongolo and Pissarides [2006].

Figure 7: Within TTWA variance components vs local LM size

(a) As share (b) As variance



Notes: We compute the variance of individual fixed effects (blue line), the variance of job fixed effects (red line), and the covariance between the two (green) line for each travel-to-work area. We then plot a binned scatterplot against log-population within the ttwa and draw quadratic lines of best fit for all three components.

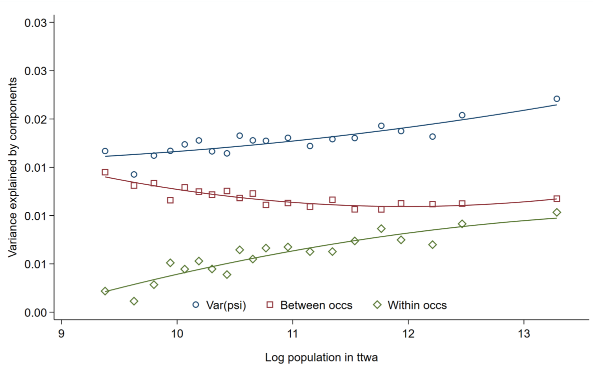
[19] In the main analysis we focus on the 286 commuting zones which have a population greater than 10,000.

[20] In appendix A.9, we show that the estimated relationships hold even after controlling for local occupation composition.

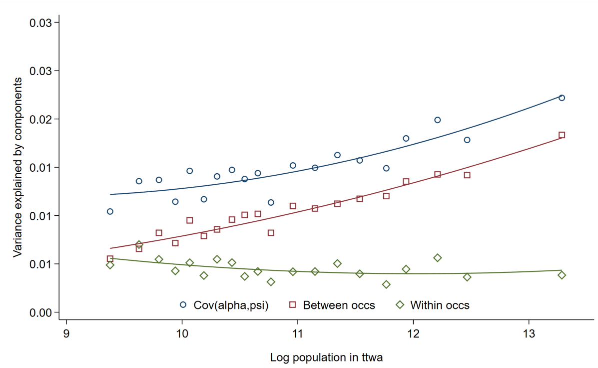
We further decompose the variance of job fixed effects and the covariance between individual and job fixed effects into their between-occupation and within-occupation components using the plug-in estimator for the between-occupation variance in each commuting zone, and relate these estimates to the size of the commuting zones. We plot these results in figure 8. Panel 8a plots the decomposition for the variance of job fixed effects, while panel 8b plots the decomposition for the sorting component. We find that the importance of within-occupation variation increases with the size of the commuting zone, while the size of between-occupation variation falls with LLM size. In smaller commuting zones, most of the variance in job FEs tend to be between occupations, with there being no variation within occupations. However, we do not find that the sorting between individuals and firms within occupations becomes more important with commuting zone size. Instead, we find that sorting of individuals between occupations is increasing in the size of the commuting zone, while the sorting of individuals within occupations seems stagnant.

Figure 8: Decomposing the within TTWA job variance and sorting components into within and between occupation elements

(a) Var Job FE



(b) Sorting



Notes: This figure presents binned scatterplots of the between- (red squares) and within-occupation (green diamonds) components of the variance of job fixed effects in panel 8a, and the covariance of individual and job fixed effects in panel 8b, against log commuting zone size. The overall size of the variance component is also plotted, in blue circles.

, firm heterogeneity matters more, but that there is not much more sorting to firms than in smaller labor markets. There are many possible explanations for this potentially surprising result. One is that as labor markets grow returns to occupation sorting (due to increasing options) increases at a rate faster than returns to firm sorting. Another possibility is that this is a feature of the French setting whereby strong labour market protection induces increased sorting on non-wage amenities in larger labor markets.

4.3 Discussion

The heterogeneity analysis performed in this section confirms the main qualitative takeaways from the overall analysis with two nuances. Overall, we find that in most labour markets occupation-level heterogeneity and sorting onto occupations is at least, or more, important than firm-level heterogeneity and sorting onto firms. However, we also find evidence that firm-level heterogeneity matters more in higher-wage occupations and larger labor markets, although the importance of sorting of individuals to firms does not increase with occupation-wage or labor market size.

Although sorting between workers and firms seems to be relatively less important in our analysis than in recent work on the sources of wage inequality, particularly the literature focused on firms (e.g. Card et al. [2013] and Song et al. [2019]). It is possible that this is due to differences in setting, and that in the French context, sorting on amenities, which is not captured in analyses based on wages, may be more important. However, it is also possible that part of the increase in sorting between workers and firms captured in this literature may be due to (1) sorting between workers and occupations and (2) the distribution of occupations in different firms. In that case, what was previously identified as worker-firm sorting may be actually worker-occupation sorting.

5 Conclusion

In this paper, we estimate the relative importance of occupation-based (e.g. heterogeneity in prices for skills and tasks) and firm-based factors (e.g. wage dispersion due to varying productivity) in contributing to overall wage inequality. To do so, we extended the classic worker-firm wage model by allowing firms to pay varying wage premia across occupations and estimate a two-way fixed effects model with worker and job fixed effects. We apply our model to the universe of French workers from 2015-19 with log annual earnings as our main wage variable, and show that our conclusions also hold in an earlier period in France (2010-14) and in Germany (1999-2022). We address concerns about limited mobility bias using the approach of Kline et al. [2020]. We also assess the robustness of our conclusions to using occupational classifications of different fineness, considering only women, and using log hourly wages instead of log annual wages.

We have three main empirical conclusions. First, we find that our decompositions paints a different picture to those suggested by worker-firm decompositions using a standard worker-firm AKM model. Relative to decomposition of the AKM model, we find a smaller role for individual heterogeneity (by 8.1 pp), a larger role for job heterogeneity (by 4.2 pp), and a larger role for sorting between workers and jobs (by 11.9 pp).

Second, we find that between-occupation job variance is roughly as important as within-occupation between-firm variation, and that between-occupation sorting is significantly more important than within-occupation between-firm sorting in our context. Our results suggest that 21% of total log-wage variance can be attributed to between-occupation pay variance and sorting of workers between occupations, while 9% can be attributed to variance within occupations between firms and sorting between workers and firms within occupations. This suggests that occupation-based factors, such as heterogeneity in skill prices, are quantitatively more important than firm pay dispersion.

We find that these average results also qualitatively hold across heterogenous labor markets with two nuances. First, in higher-wage occupations and larger labor markets firm-level heterogeneity is relatively more important, although the importance of firm-worker sorting does not vary. Second, individual-level differences are significantly more important in both higher-wage occupations and larger labor markets.

Lastly, our results suggest that firms offer different wage premia across occupations and that there is room for wage determination models to incorporate occupations to a greater extent. For example, this finding supports models which predict differing wage premia for workers in different jobs within a firm, such as Haanwinckel [2023] and Bloesch et al. [2022].

References

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, 1999.

John M. Abowd, Robert H. Creecy, and Francis Kramarz. Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau, March 2002.

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011.

Jorge Alvarez, Felipe Benguria, Niklas Engbom, and Christian Moser. Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics*, 10(1): 149–89, 2018.

M. J. Andrews, L. Gill, T. Schank, and R. Upward. High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 171(3):673–697, 03 2008. ISSN 0964-1998. doi: 10.1111/j.1467-985X.2007.00533.x. URL https://doi.org/10.1111/j.1467-985X.2007.00533.x.

Martyn J Andrews, Leonard Gill, Thorsten Schank, and Richard Upward. High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, 117(3):824–827, 2012.

David H Autor, Frank Levy, and Richard J Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333, 2003.

Miren Azkarate-Askasua and Miguel Zerecero. No more limited mobility bias: Exploring the heterogeneity of labor markets. *Working Paper*, 2024.

Damien Babet, Olivier Godechot, and Marco G Palladino. In the land of akm: Explaining the dynamics of wage inequality in france. 2022.

Justin Bloesch, Birthe Larsen, and Bledi Taska. Which workers earn more at productive firms? position specific skills and individual worker hold-up power. 2022.

Stéphane Bonhomme, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. How much should we trust estimates of firm effects and worker sorting? *Journal of Labor Economics*, 41(2):291–322, 2023.

David Card, Jörg Heining, and Patrick Kline. Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015, 2013.

David Card, Ana Rute Cardoso, and Patrick Kline. Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633–686, 2016.

David Card, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70, 2018.

David Card, Jesse Rothstein, and Moises Yi. Industry wage differentials: A firm-based approach. *Journal of Labor Economics*, 42(S1):S11–S59, 2024.

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

Deborah Goldschmidt and Johannes F Schmieder. The rise of domestic outsourcing and the evolution of the german wage structure. *The Quarterly Journal of Economics*, 132(3): 1165–1217, 2017.

Maarten Goos and Alan Manning. Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics*, 89(1):118–133, 2007.

Daniel Haanwinckel. Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution. Technical report, National Bureau of Economic Research, 2023.

Insee. Base Tous Salariés : fichiers Salariés - 2022, 2024. URL http://datapresentation.casd.eu/10.34724/CASD.137.5388.V1.

Koen Jochmans and Martin Weidner. Fixed-effect regressions on network data. *Econometrica*, 87(5):1543–1560, 2019.

Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten. Leave-out estimation of variance components. *Econometrica*, 88(5):1859–1898, 2020.

Thibaut Lamadon, Magne Mogstad, and Bradley Setzler. Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review*, 112(1):169–212, 2022.

Ilse Lindenlaub. Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, 84(2):718–789, 2017.

Jeremy Lise and Fabien Postel-Vinay. Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110(8):2328–76, 2020.

Dale Mortensen. *Wage dispersion: why are similar workers paid differently?* MIT press, 2003.

Barbara Petrongolo and Christopher Pissarides. Scale effects in markets with search. *The Economic Journal*, 116(508):21–44, 2006.

Fabien Postel-Vinay and Jean-Marc Robin. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350, 2002.

Jae Song, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. Firming up inequality. *The Quarterly journal of economics*, 134(1):1–50, 2019.

Sónia Torres, Pedro Portugal, John T Addison, and Paulo Guimaraes. 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, 2018.

A Appendix

A.1 Summary statistics for 2010-14

Table 4: Summary statistics, 2010-14

	Full data	Firms connected set	Firms LOO set	Jobs connected set	Jobs LOO set
2010-14					
N obs	46,103,347	39,467,309	37,286,700	33,154,058	28,775,013
N workers	12,664,742	9,400,612	8,890,944	7,934,805	7,034,652
N firms	1,146,870	580,256	329,961	502,751	222,969
N jobs	4,989,214	3,561,234	2,921,078	2,209,007	821,744
Mean log annual wage	10.34	10.37	10.38	10.36	10.38
Var log annual wage	0.24	0.23	0.23	0.23	0.23
Mean log hourly wage	2.89	2.91	2.92	2.91	2.92
Var log hourly wage	0.19	0.19	0.19	0.19	0.18
Var residualised log hourly wage	0.23	0.22	0.22	0.21	0.21
N moves	8,503,206	2,954,239	2,694,455	6,003,822	4,562,592
N firm moves	3,131,395	2,954,239	2,694,455	2,760,518	2,116,072
N occ moves	7,199,647	2,660,426	2,395,688	4,855,056	3,569,830
N firm + occ moves	1,827,836	1,735,055	1,550,929	1,611,752	1,123,310

Notes: This table shows the summary statistics for data from 2010 to 2014, cleaned in the same way as our main sample, which we use as a robustness check. The underlying data is from yearly BTS data files, French administrative matched employee-employer data. Individuals are mapped over time using the procedure and kindly provided programs in Babet et al. [2022]. Sample construction and restrictions are discussed in the main text.

A.2 Main decomposition robustness

A.2.1 Additional robustness checks

In table 5, we summarise the results of robustness exercises on our main log wage variance decomposition. For each exercise, we report the uncorrected fixed effects decomposition (denoted FE) and the heteroskedasticity-robust correction due to Kline et al. [2020] (denoted HE). We report the robustness of our results to (1) considering an earlier time period 2010-14, (2) the use of hourly wages instead of annual wages, and (3) considering women instead of men. We find that our results are qualitatively very similar to our baseline results in each case.

Table 5: Robustness of baseline decomposition results to alternative samples and variables

	Y10-14 FE	Y10-14 HE	Hourly wages FE	Hourly wages HE	Women FE	Women HE
Worker	0.644	0.507	0.610	0.557	0.633	0.457
Job	0.174	0.115	0.154	0.132	0.181	0.116
Sorting	0.097	0.204	0.183	0.222	0.078	0.196
Resid	0.084	0.175	0.053	0.089	0.109	0.232

Notes: This table shows the log-wage variance decomposition results for our preferred worker-job model for three alternative samples. For each exercise, we report the uncorrected fixed effects decomposition (denoted FE) and the heteroskedasticity-robust correction due to Kline et al. [2020] (denoted HE). We report the robustness of our results to (1) considering an earlier time period 2010-14 (columns 1 and 2), (2) the use of hourly wages instead of annual wages (columns 3 and 4), and (3) considering women instead of men (columns 5 and 6).

A.2.2 Decomposition on the connected set using the homoskedasticity-robust correction

To implement the Kline et al. [2020] correction, we require a leave-out connected set, which is a stronger connectedness requirement than a connected set, which is what is required for the estimation of AKM type models. Thus, the results that we present are for a selected sample of more connected workers and firms. How robust are our results to the sample selection?

The HE correction is not feasible in the connected set. Instead, in table 6, we present results for the uncorrected decomposition (FE) as well as the homoskedasticity-robust correction due to Andrews et al. [2008] (HO), for both the leave-out observation connected set (LOO) and the connected set. In column 3, we present our baseline results, for the LOO connected set using the Kline et al. [2020] corrections.

The results suggest two main conclusions: first, heteroskedasticity seems to be a feature in our sample. We can see that we get materially different decompositions using the homoskedasticity-robust correction (column 2) relative to the heteroskedasticity-robust correction (column 3). Thus, we should not expect the homoskedasticity-robust corrections in column 5 to be a good approximation of the true decomposition in the connected set.

Second, we find that in the connected set, the share due to sorting is fair lower than in the leave-out observation connected set, and the share due to individual and job heterogeneity is higher. Since firms and workers in the leave-out observation connected set are by definition more connected than those in the connected set, this result might be rationalised as due to more connected firms and workers being able to achieve more efficient sorting. Unfortunately, the true decomposition in the connected set cannot be computed as we had also found that homoskedasticity is not a reasonable assumption to make for our data.

Table 6: Baseline decomposition results

	LOO FE	LOO HO	LOO HE	Connected FE	Connected HO
Worker	0.639	0.583	0.493	0.700	0.620
Job	0.165	0.139	0.106	0.242	0.186
Sorting	0.104	0.149	0.210	-0.029	0.067
Resid	0.093	0.129	0.192	0.087	0.126

Notes: This table shows FE (uncorrected), HO (Andrews et al. [2008]) and HE (Kline et al. [2020]) decompositions for both the connected set and the leave-out observation connected set. Columns 1-3 present FE, HO and HE results for the leave-out observation connected set respectively while columns 4 and 5 present results for the FE and HO decompositions for the connected set.

A.2.3 Varying granularity of the occupation definition

In table 7, we present a robustness exercise analysing how our log wage decomposition varies with the granularity of the occupation code we use to generate firm-occupation pairs. We contrast the baseline jobs definition with 4 digit occupation codes from the PCS classification to analogous definitions with 3 digit, 2 digit and 1 digit occupation codes. In each column, we present the KSS corrected decomposition from the preferred worker-job model. We find that decreasing the granularity of occupations does marginally increase the share due to worker heterogeneity and decrease the share due to job heterogeneity and sorting. However, even with 1-digit occupations, we find that our qualitative results in contrast to the AKM model holds — we find that the share due to individual heterogeneity falls by 5.1 pp, the share due to job heterogeneity increases by 1.9 pp and the share due to sorting roughly doubles (increases by 8.9 pp).

Table 7: Robustness of baseline decomposition results to the granularity of occupation in job definition

	Occ 4D	Occ 3D	Occ 2D	Occ 1D
Worker	0.493	0.514	0.523	0.523
Job	0.106	0.097	0.091	0.083
Sorting	0.210	0.192	0.185	0.180
Resid	0.192	0.197	0.201	0.214

Notes: This table shows the robustness of the main log-wage variance decomposition results to alternative job definitions when we consider less granular occupations. In each column, we present the KSS corrected decomposition from the preferred worker-job model, where the job is defined as a firm-occupation pair where the occupation at the stated granularity is used. Column 1 presents our baseline results using 4-digit occupations and columns 2, 3 and 4 reports results with 3 digit, 2 digit and 1 digit occupations respectively.

A.2.4 Only considering cross-firm moves

In table 8, we present the results of a robustness exercise, where we drop from the sample all workers who changed jobs within a firm, and perform the same decomposition as in table 2. This is because it is possible that different kinds of moves between jobs satisfy the identification assumptions to different degrees; in particular, it may be possible that promotions and demotions within a firm may to a greater extent represent match effects that are not well-captured by individual fixed effects.

We find that the share of individual heterogeneity is substantially lower (by 10.3pp) in this alternative sample, and the share due to job FE and sorting is higher (by 2.5pp and 1.7 pp) respectively. Thus, excluding within firm occupation moves seems to emphasise the distinction between the worker-jobs model and the worker-firms model.

Table 8: Decomposition results excluding within-firm moves

	Firm AKM	Firm KSS	Job AKM	Job KSS
Worker	0.790	0.586	0.615	0.390
Firm	0.104	0.064		
Job			0.209	0.131
Sorting	0.009	0.082	0.082	0.227
Resid	0.096	0.268	0.094	0.252

Notes: This table shows the main log-wage variance decomposition results for an alternative sample where we exclude all workers who had changed occupations within the same firm. We use the largest connected leave-one-out set. As in the main table, columns one and three present a naive decomposition whereas columns two and four present a corrected decomposition due to Kline et al. [2020]. Columns one and two apply a worker-firm model and columns three and four apply a worker-job model.

A.3 Results from the German administrative data

We corroborate our main results using administrative matched employee-employer data covering the universe of individuals in Germany. We use data from the German social security records, specifically the Employee Histories (BEH). This dataset contains data on all spells of employment in Germany, excluding civil servants in the public sector. We have data from 1999 to 2022.[21] We divide this period into four overlapping time periods: 1999-2004, 2004-09, 2012-17, and 2017-22. We exclude the years 2010-11 because of the introduction of the new occupation classification system KldB 2010, which led to an implausibly large share of workers changing occupations during this period, possibly due to firms updating their registrations with the social security system.

Our main wage variable is real gross labour earnings per day, deflated to 2015 levels. One limitation of the data is that it is top-coded, and thus we lose information on the wages of the highest-earning workers in the German context. It is due to this top-coding that we focus only on the French data in the main text; while the top-coding affects only a minority of workers, we think our choices about imputation will materially affect the measures of log wage variance we find. We follow the methodology in our treatment of French data (which follows the sample harmonisation methods in Bonhomme et al. [2023]) by focusing

[21]Previous years from 1975 are available for West Germany and from 1990 are available for East Germany. We start from 1999 mainly because requesting additional years will increase data security restrictions on the project.

on men aged 25-60. We drop all observations in East Germany, following other AKM papers set in Germany (e.g. Card et al. [2013], Dauth et al. [2022]). We also focus on full-time employment spells, as recorded in the data, and do not impose the 32.5% of average earnings threshold as in Bonhomme et al. [2023] since we do not have all spells in our dataset. We follow previous work (e.g. Card et al. [2013], Dauth et al. [2022]) and impute the censored wages using the methodology proposed in Card et al. [2013]. Finally, instead of estimating education-specific age profiles as in Card et al. [2013], we follow Bonhomme et al. [2023] in first residualising wages on a cubic wage profile and year fixed effects, and perform the two-way fixed effect decomposition without controls on the residual.[22] Relative to the French occupational classification, the German occupational classification at the 5-digit level is more fine, with over 1000 categories, relative to the 430 categories in the French occupation definition.

A.3.1 Descriptive statistics

Table 9 shows descriptive statistics on the sample size and number of workers and firms/ jobs in the German data in total, as well as summary statistics from the largest connected set and the leave-one-out connected set of the worker-job model. Compared to the French data (described in table 1 in the main text), our German sample is of similar size. In the German data the largest leave-one-out connected set covers 60% of all observations, 56% of workers, and 19% of jobs. Comparably in the French data the largest leave-one-out connected set covers 66% of all observations, 57% of workers, and 18% of jobs, implying that the French data is about as connected than the German data.

We find that 10.8% of all observations in the entire dataset involve moves, of which 50.8% involve moves between firms within occupations, 11.8% involve moves within firms between occupations, and the remaining 37.3% involve moves between both firm and occupation. This differs from the French setting where most moves involve moves across occupations. This could be because moves across occupations within firms might be underreported in the data, as firms may not always be incentivised to update their registration with the social

[22] While modelling the age profile more flexibly is more realistic, we use the latter method to make these results more comparable to the French context, which lacks information on workers' education, and to Bonhomme et al. [2023], which reports estimates from a wide set of countries.

security authority when their workers change occupations.

A.3.2 Main decomposition

Table 10 shows the results from performing our main decomposition exercise in the German data on the largest leave-one-out connected set using the variance correction methodology of Kline et al. [2020]. In the German setting, more of the log-wage variance can be attributed to job fixed effects and sorting than in the French setting, while the share of worker fixed effects is similar in both contexts. We find that 18% of log wage variance can be explained by heterogeneity in jobs in the German context relative to 11% in France, while 21% can be explained by sorting in Germany relative to also 21% in France. Qualitatively, we also find that estimating a job fixed effects model instead of a firm fixed effects model increases the share due to sorting by about 7 pp and the share due to job FE by 8 pp.

Table 10: Log wage decomposition in West Germany, 2017-22

	Firm FE	Firm KSS	Job FE	Job KSS
Worker	0.625	0.525	0.525	0.416
Firm	0.129	0.107		
Job			0.240	0.184
Sorting	0.102	0.143	0.110	0.213
Residual	0.083	0.166	0.076	0.138
Controls	0.060	0.060	0.049	0.049

Notes: This table replicates figure 2 in the main text using data from West Germany in the period 2017-22. Column 1 produces the decomposition with the plug-in approach for the worker-firm model, while column 2 shows the KSS-corrected results for the worker-firm model. Column 3 shows the results for the worker-job model using the plug-in fixed effects approach while column 4, our preferred specification, corrects these estimates using the KSS approach.

Does the log-wage decomposition vary over time? In table 11, we plot our preferred KSS-corrected log-wage decomposition for the worker-job model in four periods from 1999 to 2022. We find that the share of log-wage variance due to individual and job fixed effects is fairly constant over time, but we find that the share due to sorting has increased from 12.3% in 1999-2004 to 21.3% by the end of the period in 2017-22%. This coheres with previous work in Card et al. [2013] and Song et al. [2019] emphasising the growing importance of sorting for log-wage inequality.

Table 9: Descriptive statistics for entire West German sample, in four periods from 1999-2022

	1999-2004	2004-09	2012-17	2017-22
Total observations (m)	60.34	57.78	62.47	54.56
Total workers (m)	13.62	12.81	14.03	14.22
Total firms (m)	1.56	1.49	1.40	1.35
Total jobs (m)	4.03	3.78	4.79	4.91
Average daily wages	4.77	4.74	4.75	4.78
SD daily wages	0.49	0.53	0.53	0.52
SD resid. daily wages	0.48	0.51	0.51	0.50
N moves	6.51	5.34	6.16	5.79
N firm moves	5.74	4.76	5.37	4.88
N occ moves	3.20	2.58	3.85	3.77
N firm + occ moves	2.43	2.00	3.06	2.86
Largest connected set				
Total observations (m)	48.69	44.81	47.24	40.16
Total workers (m)	9.89	8.98	9.56	9.44
Total firms (m)	0.99	0.87	0.85	0.78
Total jobs (m)	2.24	1.91	2.41	2.30
Average daily wages	4.80	4.78	4.77	4.80
SD daily wages	0.47	0.51	0.52	0.51
SD resid. daily wages	0.46	0.49	0.50	0.49
N moves	5.65	4.62	5.37	5.00
N firm moves	5.08	4.18	4.71	4.27
N occ moves	2.79	2.26	3.42	3.29
N firm + occ moves	2.23	1.82	2.76	2.56
Largest leave-out observation connected set				
Total observations (m)	42.22	38.06	39.32	32.96
Total workers (m)	8.73	7.76	8.17	7.95
Total firms (m)	0.54	0.45	0.44	0.39
Total jobs (m)	1.05	0.86	1.00	0.92
Average daily wages	4.81	4.81	4.79	4.82
SD daily wages	0.46	0.50	0.52	0.51
SD resid. daily wages	0.45	0.49	0.50	0.49
N moves	4.43	3.53	3.94	3.58
N firm moves	3.98	3.18	3.40	3.01
N occ moves	2.02	1.58	2.37	2.21
N firm + occ moves	1.57	1.22	1.82	1.64

Notes: This table presents summary statistics for all four periods for three samples - the entire population, the largest connected set of the main worker-job specification, and the largest leave-out observation set of the main worker-job specification. We present the total number of observations, the total number of unique workers, the total number of unique firms and the total number of unique jobs in millions, the total number of moves, as well as the average and standard deviation of the daily (imputed) wage as well as the standard deviation of the daily wage residualised on year fixed effects and a cubic age profile.

Table 11: Log wage decomposition in West Germany over time, 1999-2022

	1999-2004	2004-09	2012-17	2017-22
Share var(alpha)	0.417	0.396	0.381	0.416
Share var(psi)	0.218	0.233	0.220	0.184
Share cov(alpha,psi)	0.123	0.157	0.200	0.213
Share eps	0.187	0.154	0.140	0.138
Share controls	0.055	0.060	0.058	0.049

Notes: This table presents the log-wage decomposition using our preferred worker-job fixed effect model with KSS corrections using data from West Germany focusing on four periods between 1999-2022. The years 2010 and 2011 are excluded due to a change in the occupational classification codes during that time.

A.3.3 Robustness exercises

While Kline et al. [2020] improve on the corrections for limited mobility bias introduced in Andrews et al. [2008] by allowing for heteroscedasticity, a limitation is that it is only feasible on leave-out connected sets, which is a subset of the largest connected sets used for the estimation of the two-way fixed effect models. This may lead to a relatively selected sub-sample which may affect the results. To assess this issue, in table 12, we summarise the log-wage decomposition in both the largest connected set, and the largest leave-out observation connected set. We include the uncorrected plug-in estimates (FE), the correction from Andrews et al. [2008] under the assumption of homoscedasticity (HO), and the Kline et al. [2020] correction allowing for heteroscedasticity (HE) only possible in the leave-out observation connected set.

Table 12: Log wage decomposition in West Germany, different limited mobility corrections by sample

	fe	ho	fe	ho	he
Share var(α)	0.596	0.521	0.525	0.473	0.416
Share var(ψ)	0.333	0.275	0.24	0.21	0.184
Share cov(α, ψ)	-0.055	0.05	0.11	0.163	0.213
Share ε	0.07	0.099	0.076	0.104	0.138
Share controls	0.055	0.055	0.049	0.049	0.049
Sample	Connected	Connected	LOO	LOO	LOO

Notes: This table presents the log-wage decomposition using our preferred worker-job fixed effect model using data from West Germany in 2017-22. Columns 1 and 2 present FE and HO models using the largest connected set while columns 3-5 present FE, HO and HE (main specification) corrections within the leave-out observation connected set.

We draw two main conclusions from this exercise. First, as we find substantially different decompositions with the HO and the HE corrections, we conclude that our data seems to not be characterised by homoskedasticity. We find that the HE correction still overstates the share due to individual heterogeneity and underestimates the role of job heterogeneity as well as sorting. Second, we find that both the FE and HO corrections are substantially different in the largest connected set and the largest leave-out observation connected set. In particular, the share due to sorting seems to be much lower in the largest connected set than in the leave-out connected set. This may imply that our results overstate the extent of sorting in the whole sample, and that less connected individuals and jobs may be less well-sorted than more connected individuals and jobs.

Another robustness exercise we conduct is to analyse how the decomposition varies when occupation codes with different fineness are used to create the firm-occupation interaction effects in the main jobs model. In table 13, we present the results of varying the construction of the jobs indicator. In our baseline specification, we use German 5 digit KldB codes, where the fourth and fifth digits provide additional information on the worker's role in the company and the skill level of his job. We assess robustness to using 4 digit and 3 digit codes instead of 5 digit codes to create the job indicator variable. We find that using less fine occupation codes increases the share due to worker fixed effects and decreases the share due to sorting and job fixed effects. We also find that the share due to the residual also increases. Thus, it seems that using less fine occupation codes will lead to overstating the relative importance of individuals and understating the relative importance of jobs and sorting in log-wage variance.

Table 13: Log wage decomposition in West Germany with different job definitions

	KldB 5 digit	KldB 4 digit	KldB 3 digit	Firm
Worker	0.416	0.445	0.471	0.525
Job	0.184	0.171	0.152	0.107
Sorting	0.213	0.191	0.175	0.143
Residual	0.138	0.142	0.147	0.166
Controls	0.049	0.052	0.055	0.060

Notes: This table presents the log-wage decomposition using our preferred worker-job fixed effect model with KSS corrections using data from West Germany in 2017-22. Column 1 presents our preferred specification with job fixed effects defined as the interaction between establishments and five-digit KldB 2010 occupations. Column 4 presents estimates from the AKM model where establishments are not interacted with occupations. Column 2 presents estimates with job fixed effects defined as the interaction between establishment and KldB 2010 4 digit occupations, and column 3 presents estimates with job fixed effects defined as the interaction between establishment and KldB 2010 3 digit occupations.

A.3.4 Between-within occupation decompositions

We also reproduce the within-between occupation analysis in section 3.2 here for the German data. This is presented in table 14. We find that within-occupation job variance accounts for about 2/3 of total job FE variance, and that within-occupation sorting accounts for basically none of total sorting between individuals and firms. In comparison to the French setting, we find more of wage variation is due to between-occupation variance, and sorting between occupations instead of within occupations between firms.

A major caveat with this analyses is that the top-coding of the German data may drive these decompositions. The exact imputation procedure may also play an outsized role in the between- and within-occupation breakdown. For these reasons, we prefer the results from the French data and leave the German results as suggestive evidence showing the robustness of our qualitative results.

Table 14: Between- and within- occupation decompositions of job FE variance and indiv- and job- covariance

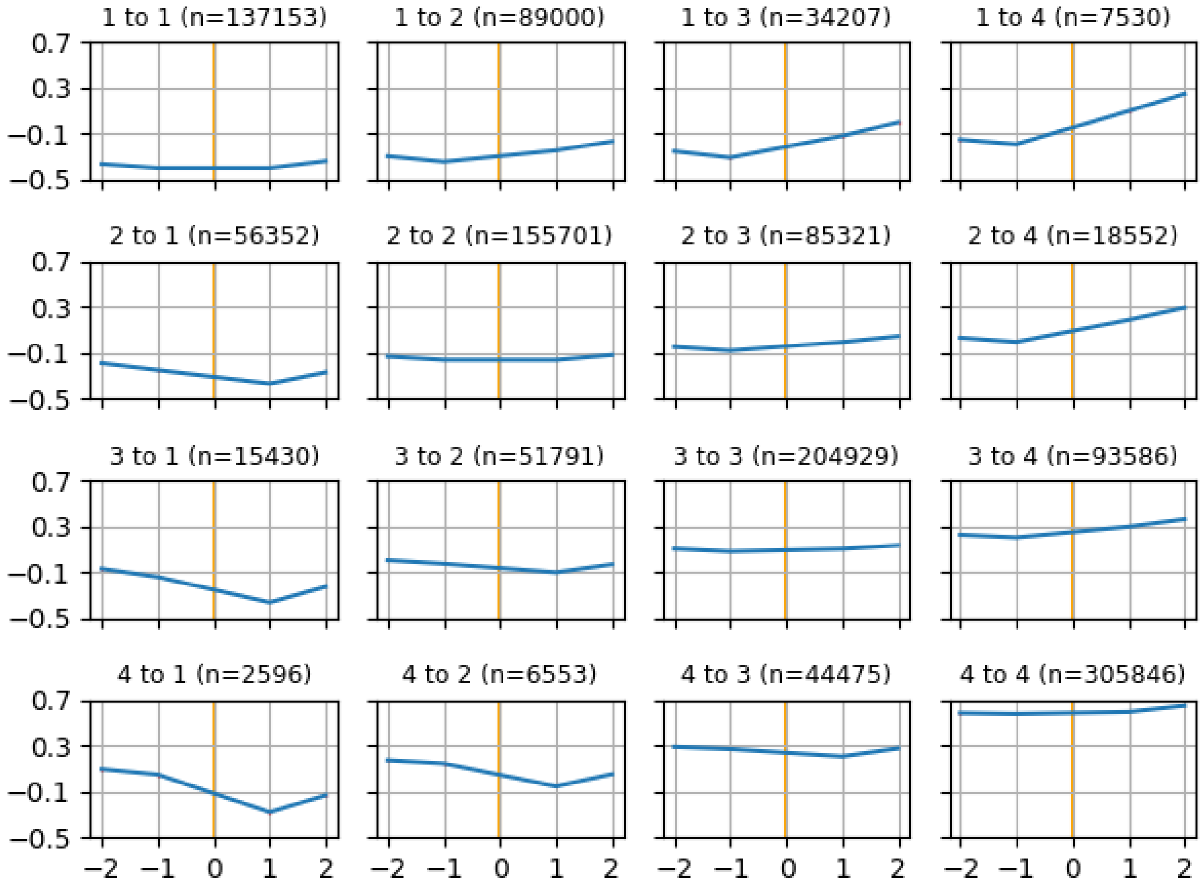
	1999-2004	2004-09	2012-17	2017-22
$Var(\psi)$	0.047	0.058	0.059	0.047
.Within	0.033	0.039	0.036	0.030
.Within (share).	0.698	0.664	0.608	0.643
.Betw	0.014	0.020	0.023	0.017
.Betw (share).	0.302	0.336	0.392	0.357
$Cov(\alpha, \psi)$	0.013	0.020	0.027	0.027
.Within	-0.001	0.001	0.001	0.001
.Within (share)	-0.043	0.038	0.030	0.046
.Betw	0.014	0.019	0.026	0.026
.Betw (share)	1.043	0.962	0.970	0.954
Var(Y)	0.214	0.251	0.267	0.255
$Var(Y_{resid})$	0.203	0.236	0.251	0.243

Notes: This table presents the within-occupation share of job FE variance and individual and job FE covariance, calculated by taking the mean of HE-corrected variance of job FE and covariance of individual and job FEs within occupations. The between-occupation variance is calculated as the residual.

A.4 Additional diagnostics testing the identifying assumption

Figure 9 plots the effect on average wages of job changes, where we aggregate jobs into four groups by the average wage within the job. These diagrams are sometimes preferred because it does not use job fixed effects, which require prior estimation of the model. We nevertheless see the expected wage changes in moving between different job quartiles, consistent with our worker-job fixed effects model. However, we also see that the average wage within the starting quartile and the ending quartile is not consistent, reflecting a degree of sorting of higher wage workers into more highly paid jobs.

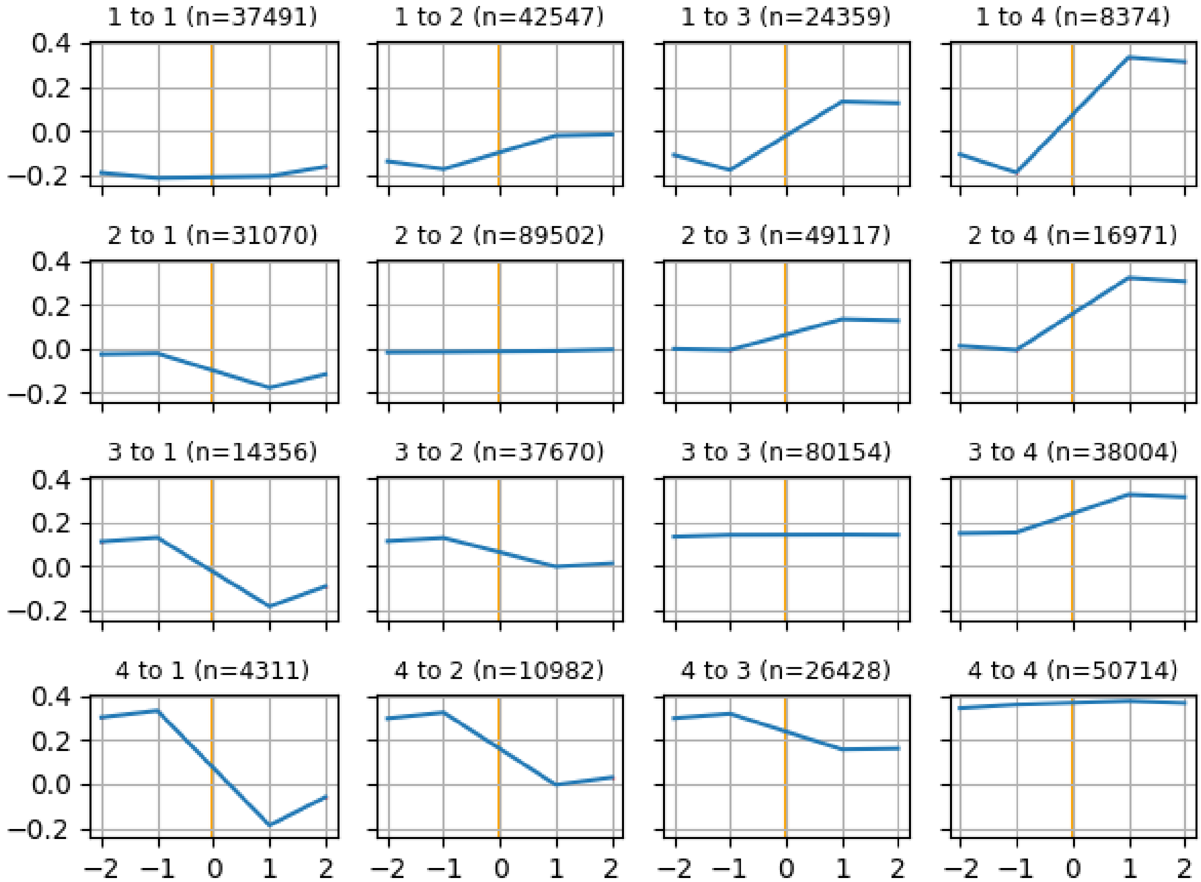
Figure 9: Event study around job moves, clustering by average wage within the job



Notes: This figure shows the impact on average wages around the event of job movement. Each cell shows a movement event from quartile to quartile of the average wage within the job. Only those who remain in their old job for two years before and their new jobs for two years after the move event are included. The number of switchers in each cell is given in the cell title.

Figure 10 plots the equivalent of figure 1 in the main text, including only moves across occupations within firms. We do this to check if the identification assumption is also valid for moves within firms across occupations, which one might suspect to behave differently as promotion events might indicate particular match effects between the worker and the job. We find that event studies around such moves within firms exhibits a stepwise structure as was in the case of including all moves, consistent with what is implied by our worker-jobs model.

Figure 10: Event study around job moves, including only moves across occupations within firms, clustering by estimated job fixed effect

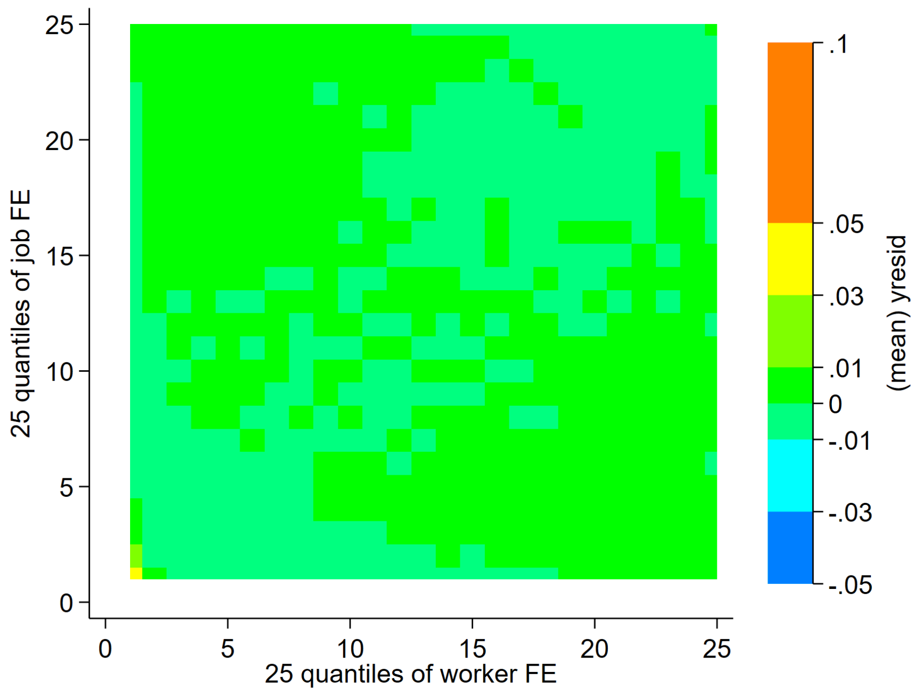


Notes: This figure shows the impact on average wages around the event of job movement between occupations within firms. Each cell shows a movement event from quartile to quartile of the average estimated job fixed effect distribution. Only those who remain in their old job for two years before and their new jobs for two years after the move event are included. The number of switchers in each cell is given in the cell title.

A.5 Is the log-additive functional form a good specification for wages?

In this section, we check whether the log additive functional form in fixed effects is a good approximation to the data. To do this, we follow the test in Card et al. [2013] which they use the argue for the log-wage model, and aggregate our estimated fixed effects into 50 quantile groups. For each worker quantile and job quantile, we take the mean of the residual of the two-way fixed effect regression, reasoning that if the specification is a good fit to the data, then the mean of the residuals should be 0 or close to zero. We plot a heat map in figure 11 with the quantiles of worker FE on the x-axis, the quantiles of job FE on the x-axis, and the size of the mean residuals as the colour within the cell. Warmer colours (yellow, orange, red) mean that the model underpredicts the true wage in the data while cooler colours (cyan, blue) means that the model overpredicts the actual wage data. Green implies that the mean residual in the cell is within 1% on each side of 0.

Figure 11: Mean of residuals by worker FE and job FE quantiles (25 groups each)



Notes: This figure plots the mean of the two-way fixed effect residuals of the main two-way fixed effect regression within each combination of 25 worker FE quantiles and 25 job FE quantile cells. A large mean residual suggests that the model is a poor fit for the data in those particular job and individual quantiles.

The figure shows that for the vast majority of cells the difference between the predicted value and the true data is within approximately 1% of the actual log wage, and for all cells, the difference is within approximately 3% of the actual log wage. Thus, it appears that the log-additive functional form in the worker-jobs model seems to fit the data well.

A.6 Specifying a worker-firm model if the true model is log-additive in worker- and job-fixed effects

In the paper, we noted that when the true data-generating process underlying the data involves worker-job (firm-occupation pair) fixed effects but a worker-firm FE two-way fixed effect regression is specified, the model leads to the overestimation of variance attributed to individuals under certain conditions. We show this in more detail in this section. This result is conceptually similar to the work in Abowd et al. [1999] to show the effects of misspecifying the model with worker-industry FE when the true model is worker-firm. This also relates to the biases due to so-called 'hierarchy effects', discussed in Card et al. [2024].

We can rewrite the true data-generating process as follows.

$$
\begin{aligned}
\ln(w_{it}) &= \alpha_i + \phi_{F(i,t)} + (\lambda_{J(i,t)} - \phi_{F(i,t)}) + \varepsilon_{it} \\
\phi_f &= \frac{1}{n_f} \sum_{j \in \bar{f}} \lambda_j
\end{aligned}
$$

It seems from this rewriting that what misspecification does is to add a term to the part that varies across individuals and periods which is now an unobserved omitted variable, and which will therefore cause the regression to estimate the individual fixed effects and the mean firm fixed effects with bias. Let W denote a $NT \times 1$ vector of wages, L a $NT \times N$ matrix of indicators for individuals $i = 1, \cdots, N$, and J a $NT \times H$ matrix of indicators for jobs $j = 1, \cdots, H$. Let F denote a $NT \times G$ matrix of indicators for the firm the job is at (for $f = 1, \cdots, G$). Let ε denote a $NT \times 1$ vector of ε_{it} errors. Finally, let K be a $H \times G$ matrix where each row $j = 1, \cdots, H$ indexes which firm the job is at. Let α denote a $N \times 1$ vector of personal fixed effects, λ a $H \times 1$ vector of job fixed effects and Φ a $G \times 1$ vector of firm-specific mean job fixed effects.

$$
\begin{aligned}
W &= L\alpha + J\lambda + \varepsilon \\
&= L\alpha + F\Phi + J(\lambda - K\Phi) + \varepsilon
\end{aligned}
$$

When J is not observed, estimating an OLS regression on L and F leads to omitted variable bias. Suppose we estimate the misspecified model. Then the parameters estimated have omitted variable bias as follows, where $M_A = I - A(A'A)^{-1}A'$ for an arbitrary matrix A.

$$\alpha^* = \alpha + (L'M_F L)^{-1} L'M_F J(\lambda - K\Phi) \tag{6}$$

$$= \alpha + (L'M_F L)^{-1} L'M_F J\lambda \tag{7}$$

$$\Phi^* = \Phi + (F'M_L F)^{-1} F'M_L J(\lambda - K\Phi) \tag{8}$$

$$= (F'M_L F)^{-1} F'M_L J\lambda \tag{9}$$

The M_F and M_L matrices are residualising matrices, where $M_F L$ refers to the residual of a linear projection of F on L as in the Frisch-Waugh-Lovell theorem. It is instructive to re-write the expressions above as follows.

$$\alpha^* = \alpha + \beta \tag{10}$$

$$\Phi^* = \iota \tag{11}$$

Where β is the regression coefficient found from estimating the following by OLS: $J\lambda = (M_F L)\beta + \xi$ and ι is similarly the regression coefficient from estimating $J\lambda = (M_L F)\iota + \xi$. Thus, the bias in variance estimation is easily characterisable.

$$\mathbb{V}[\alpha^*] - \mathbb{V}[\alpha] = \mathbb{V}[\beta] + 2Cov(\alpha, \beta) \tag{12}$$

$$\mathbb{V}[\Phi^*] - \mathbb{V}[\Phi] = \mathbb{V}[\iota] - \mathbb{V}[\Phi] \tag{13}$$

From this formula we can see that the individual fixed effects will be over estimated if $Cov(\alpha, \beta)$ is not sufficiently negative. To gain intuition as to the signs and magnitudes of $\mathbb{V}[\beta]$ and $2Cov(\alpha, \beta)$, we must understand β. β is an $N \times 1$ vector of coefficents, an individual element of β is large if for that individual the correlation between the firm-residualed individual and job effect is large. If the strength of this relationship varies over individuals $\mathbb{V}[\beta]$ will be large. If this relationship is stronger for higher-FE individuals $Cov(\alpha, \beta)$ will be positive, this is the most likely scenario.

It's harder to sign the difference in firm fixed effect variance. ι is the firm-analagous version of β, it is large if the firm conditional on worker is correlated with job FE. However, it's unclear whether the strength of variation in this coefficient will be less than or greater than the variation in firm effects itself and so signing the difference is difficult.

In general, from the above we can say that we expect the variance of the estimated worker and firm fixed effects in the mis-specified model to be wrong, and for it to be increasingly wrong the more residual variation there is in the relationship over workers between workers and jobs (conditional on firms) and over firms between firms and jobs (conditional on workers).

A.7 Comparing log-wage decomposition results against other results in the literature

In this section we compare results from estimating the usual worker-firm AKM model decomposition on our UK data to that from some previous studies using data in other countries. Table 15 summarises a few previous results[23] and compares it to our estimates using the French and German data. Panel [A] compares results using the basic AKM decomposition which does not correct for limited mobility bias whereas panel [B] shows results correcting for limited mobility bias using Kline et al. [2020]'s heteroskedasticity-robust bias correction. The main take away from this table is that in these specifications, our results appears to be in the ballpark of comparable estimates from other contexts.

[23] Bonhomme et al. [2023] provides a more exhaustive summary of the literature. We provide this table to contextualise our results in the context of the previous literature.

Table 15: Select naive and bias-corrected worker-firm decompositions from previous studies.

[A] Basic AKM Decomposition					
Paper	**Period**	**Country**	**WFE**	**FFE**	**Sort**
Card et al. [2013]	2002-09	Germany	51%	21%	16%
Card et al. [2016]	2002-09	Portugal	58%	20%	11%
Song et al. [2019]	2007-13	USA	52%	9%	12%
Alvarez et al. [2018]	2008-12	Brazil	58%	16%	18%
Our data	2015-19	France	78%	10%	2%
Our data	2017-22	Germany	63%	13%	10%
[B] Bias Corrected results					
Paper	**Period**	**Country**	**WFE**	**FFE**	**Sort**
Babet et al. [2022][1]	2014-19	France		6%	13%
Bonhomme et al. [2023]	2010-15	US		6%	13%
Bonhomme et al. [2023]	2010-15	Austria		13%	13%
Bonhomme et al. [2023]	1996-2001	Veneto, Italy		16%	11%
Bonhomme et al. [2023]	2009-14	Norway		12%	13%
Bonhomme et al. [2023]	2002-05	Sweden		7%	5%
Our data	2015-19	France	57%	6%	9%
Our data	2017-22	Germany	53%	11%	14%

Notes: This table compares results using the German and French data used in this paper when estimating classic and heteroskedasticity-robust bias-corrected AKM models, to those found in selected papers from the previous literature. Some papers omit reporting the worker-component of the log-wage decomposition. (1) Babet et al. [2022] uses an alternative split-sampling methodology described in their paper.

A.8 Computing KSS corrections to conditional variances/covariances

A.8.1 AKM variances in matrix notation

We begin by following the worked examples in Kline et al. [2020] applied to the AKM model without controls. Let D be the $n \times N$ matrix of personal indicators, and F be the $n \times J$ matrix of firm indicators.

Let $X = \begin{bmatrix} D & F \end{bmatrix}$ denote the covariates, a $n \times (N + J)$ matrix. Let $\beta = \begin{bmatrix} \alpha \\ \lambda \end{bmatrix}$ denote a $(N + J) \times 1$ vector of coefficients comprising of sub-vectors α, a $N \times 1$ vector of individual FEs and λ, a $J \times 1$ vector of firms FEs.

Suppose we want to estimate e.g. the variance of the individual fixed effects. To get this, we note that an employment weighted vector of the individual fixed effects (which is what we usually calculate) is given as:

$$\tilde{\alpha} = \begin{bmatrix} D & 0 \end{bmatrix} \beta$$

The employment weighted vector of firm fixed effects is given as:

$$\tilde{\lambda} = \begin{bmatrix} 0 & F \end{bmatrix} \beta$$

To express the variances in the matrix form $\beta' A \beta$, we get the following expression for e.g. the variance of the firm fixed effects, as follows, where we denote $\Omega_n \equiv \left(I - \frac{1}{n} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \right)$, which is a square matrix with dimensions $n \times n$.

$$Var(\tilde{\lambda}) = \frac{1}{n} \beta' \begin{bmatrix} 0 \\ F \end{bmatrix} \Omega_n' \Omega_n \begin{bmatrix} 0 & F \end{bmatrix} \beta$$

Then, the A corresponding to $Var(\tilde{\lambda})$, denoted $A_{\tilde{\lambda}}$, is:

$$A_{\tilde{\lambda}} = \frac{1}{n} \begin{bmatrix} 0 \\ F \end{bmatrix} \Omega_n' \Omega_n \begin{bmatrix} 0 & F \end{bmatrix}$$

By similar reasoning, the A corresponding to $Var(\tilde{\alpha})$, denoted $A_{\tilde{\alpha}}$ is:

$$A_{\tilde{\alpha}} = \frac{1}{n} \begin{bmatrix} D \\ 0 \end{bmatrix} \Omega_n' \Omega_n \begin{bmatrix} D & 0 \end{bmatrix}$$

Similarly, the Q corresponding to the covariance of $\tilde{\alpha}$ and $\tilde{\lambda}$, denoted $A_{\tilde{\alpha},\tilde{\lambda}}$ is:

$$A_{\tilde{\alpha},\tilde{\lambda}} = \frac{1}{n} \begin{bmatrix} D \\ 0 \end{bmatrix} \Omega_n' \Omega_n \begin{bmatrix} 0 & F \end{bmatrix}$$

A.8.2 Conditional variances in matrix notation

Now suppose that there is now a $n \times 1$ vector denoting membership into a group, G, populated with 1s and 0s. Denote this by G. Then, we note that $n_g \equiv \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} G$.

A simple way to proceed is as follows. First, we use a selection matrix E to select only the elements that are in the group, resulting in a new vector that includes only these elements.

$$\underbrace{E_G}_{\text{Selector matrix corresponding to vector G, dimensions:}n_g \times n} \overbrace{X}^{\text{Original } n\times 1 \text{ matrix}} = \underbrace{X_G}_{n_g \times 1 \text{ vector of selected elements}}$$

To construct E_G, first create an identity matrix with G as its trace IG. Then, if the ith element of G is 0, remove the ith row of IG. The resulting matrix is the $n_g \times n$ matrix, E_G.

The variance of X_G is thus:

$$Var(X_G) = \frac{1}{n_g} X' E_G' \Omega_{n_g}' \Omega_{n_g} E_G X$$

A.8.3 Conditional variances and covariances in AKM

We apply the concept from the previous subsection to the representing the conditional variances and covariances of worker and firm effects in the quadratic form. An alternative, bootstrap-based, approach to calculating condition variances is described by Azkarate-Askasua and Zerecero [2024].

Let G again denote a vector of 1s and 0s denoting membership in the target group. Then denote the conditional employment weighted vector of worker FEs by $E_G\tilde{\alpha}$ and the conditional employment weighted vector of firm FEs by $E_G\tilde{\lambda}$. Both these vectors are of length n_G.

Then, the A corresponding to $Var(E_G\tilde{\lambda})$, denoted $A_{E_G\tilde{\lambda}}$, is:

$$A_{E_G\tilde{\lambda}} = \frac{1}{n_G} \begin{bmatrix} 0 \\ F \end{bmatrix} E_G' \Omega_{n_G}' \Omega_{n_G} E_G \begin{bmatrix} 0 & F \end{bmatrix}$$

By similar reasoning, the A corresponding to $Var(E_G\tilde{\alpha})$, denoted $A_{E_G\tilde{\alpha}}$ is:

$$A_{E_G\tilde{\alpha}} = \frac{1}{n_G}\begin{bmatrix} D \\ 0 \end{bmatrix} E_G'\Omega_{n_G}'\Omega_{n_G}E_G\begin{bmatrix} D & 0 \end{bmatrix}$$

Similarly, the A corresponding to the covariance of $\tilde{\alpha}$ and $\tilde{\lambda}$, denoted $A_{\tilde{\alpha},\tilde{\lambda}}$ is:

$$A_{\tilde{\alpha},\tilde{\lambda}} = \frac{1}{n_G}\begin{bmatrix} D \\ 0 \end{bmatrix} E_G'\Omega_{n_G}'\Omega_{n_G}E_G\begin{bmatrix} 0 & F \end{bmatrix}$$

We then recover the conditional variances by plugging in the relevant A matrix in equation 1 of Kline et al. [2020]. For computation, we modify the PyTwoWay package [Bonhomme et al., 2023] to allow conditional variances by supplying custom values of the A matrix.

A.9 Is the relation between the size of travel to work area and variance due to occupational composition?

The cross-geography patterns described in figure 7 could be a reflection of differing occupation compositions across space, rather than any geography-specific factors. To consider this possibility, we can control for geographic occupation composition by conditioning on the proportion employed in each of the nine 1-digit industries in each geography.

Table 16 shows the results from this exercise. It compares the cross-CZ relationship between the CZ-specific share of wage variance due to workers, jobs, and sorting and log CZ size controlling or not for CZ occupation composition at the one-digit level. Odd columns are unconditional and reflect the results in the main text, and even columns are conditional on one-digit occupation composition. Comparing column pairs one can see that compositional effects are not driving the results.

Table 16: Controlling for geographic composition

	(1) Worker	(2) Worker	(3) Job	(4) Job	(5) Sorting	(6) Sorting
Log CZ pop	1.094***	1.003***	-0.459***	-0.463***	0.866***	0.859***
	(0.165)	(0.169)	(0.173)	(0.178)	(0.309)	(0.318)
Composition		Yes		Yes		Yes
r2	0.134	0.165	0.0241	0.0597	0.0270	0.0553
N	286	286	286	286	286	286

Notes: This table shows the results from regressing log population against wage variance decomposition shares at the commuter zone level. Columns one and two show the results with worker share, column three and four with job share and columns five and six with the share due to sorting. Odd columns show unconditional results whereas even columns condition on the proportion employed in each of the nine one-digit industries (omitting one industry due to multicollinearity concerns). Standard errors are robust. Stars indicate the usual significance levels. Commuter zones with a population below 10,000 have been omitted.