Do Workers Sort to Firms or to Occupations?

Luke Heath Milsom* Shihang Hou*

October 2, 2025

Abstract

In this paper, we find that high-wage workers sort mainly to high-wage occupations and not to high-wage firms, and that half of the previously documented sorting to firms can be attributed to the segregation of occupations across firms. To reach these conclusions, we leverage the universe of matched employee-employer data from France and Germany and estimate a flexible two-way worker-job fixed effects model of log wages. We then isolate worker sorting to firms by studying the within-occupation across-firm covariance between worker and job fixed effects.

*Heath Milsom: Institute for Fiscal Studies and KU Leuven, luke.heathmilsom@kuleuven.be. 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).

Two central facts in the study of wage inequality are that some firms pay persistently more than others [Abowd et al., 1999, Mortensen, 2003, Card et al., 2013, 2016, Song et al., 2019], and some occupations command higher pay than others [Autor et al., 2003, Goos and Manning, 2007, Acemoglu and Autor, 2011]. In this paper, we argue that it is crucial to consider both firms and occupations in conjunction because high-paying occupations are unevenly distributed across firms. This leads to the risk that focusing solely on firms conflates returns from sorting into high-paying firms with the returns from sorting into high-paying occupations.[1] The potential for confounding is empirically important; in our German and French administrative data, we find that occupations are highly segregated across firms.

To disentangle firm and occupation sources of wage determination and sorting, we adapt the two-way fixed effects framework of Abowd et al. [1999] (AKM). We begin by defining a "job" as a firm-occupation pair, and estimate a two-way model recovering worker and job fixed effects. We then use the Law of Total Covariance to decompose the covariance of worker-job fixed effects into two separate components: sorting of workers to firms within an occupation, and sorting of workers to occupations. This decomposition separates worker-job sorting into two elements, each relevant to a different theoretical literature: the first component isolates the firm-level sorting that is central to search and matching models of the labour market [Sattinger, 1993, Postel-Vinay and Robin, 2002, Shimer and Smith, 2000], while the second component captures the occupational sorting central to e.g. Roy-style models of self-selection[Heckman and Sedlacek, 1985, Keane and Wolpin, 1997]. We implement this decomposition using French and German administrative data, applying the leave-one-out method of Kline et al. [2020] to correct for biases from limited worker mobility [Andrews et al., 2008, Bonhomme et al., 2023], and checking for exogeneity using the event-study approach of Card et al. [2013].

In both the French and German settings, we find that sorting to occupations is quantitatively over four times as important as sorting to firms, and that over half of the sorting to firms found in a standard AKM decomposition can be explained by the clustering of occupations across firms. Drawing on our French data, we find that while the canonical model attributes 9.1% of wage variance[2] to worker-firm sorting, our decomposition finds that sorting across firms within occupations accounts for only 4.0%. In contrast, the sorting of workers to occupations accounts for nearly 17.0% of the variance. Our model also reveals that the total importance of workplace premia is larger than previously understood. We find that within-occupation firm heterogeneity accounts for 5.0% of wage variance, a figure

[1] For example, computer scientists may join Google not because it offers a generous firm-wide pay premium to all its workers, but because it is one of the few employers that offers elite software engineering roles. Even if Google paid industry-average wages in these high-value occupations, a researcher using the firm-premium approach would estimate a large Google premium, simply due to its concentration of high-value occupations.

[2] We present all results relative to residualised log wage variance, where we control for a cubic age profile and year fixed effects.

comparable to the 6.4% estimate for the entire firm effect in the standard model. However, we find that differences in pay premia between occupations account for a further 6.0% of total log wage variance, bringing the total share of log wage variance explained by job fixed effects to 11.0%. Our findings are quantitatively similar in the German context and are robust to using coarser occupational classifications.

Our work primarily contributes to the large literature that combines theoretical models of assortative matching [Sattinger, 1993, Postel-Vinay and Robin, 2002, Shimer, 2005] with empirical work using the two-way fixed effects model of Abowd et al. [1999], Card et al. [2013], Song et al. [2019]. We show that when high-skill occupations cluster within firms—a prevalent feature of the data—standard estimates of worker-firm sorting are confounded by worker-occupation sorting. By separating these channels, we provide the first flexible estimates of the relative importance of worker-firm sorting (e.g., matching on productivity) and worker-occupation sorting (e.g., Roy-style selection) [Heckman and Sedlacek, 1985, Lindenlaub and Postel-Vinay, 2023].

Our paper also contributes to the literature that explores the role of occupations and firms jointly in wage determination using an AKM approach. While prior work has sought to incorporate occupations, it has relied on more parsimonious modelling assumptions such as log-additive separability [Torres et al., 2018] or interactions with only broad occupational categories [Goldschmidt and Schmieder, 2017, Lamadon et al., 2022]. In contrast, our job-fixed-effect model allows for the job premium to vary flexibly, unrestricted by firm or occupation main effects. This flexibility is crucial, as imposing log-additivity may mechanically understate the variability of wage premia across firms within occupations.

The paper proceeds as follows. Section 1 presents the econometric approach. Section 2 describes the French and German administrative data used in our decomposition. Section 3 presents the main decomposition results, robustness, and validation exercises. Finally, section 4 concludes.

1 Separately identifying worker-firm and worker-occupation wage-sorting

We consider a framework that allows us to identify sorting between workers and firms, taking occupations into account in the most flexible way we can. To do this, we augment the classic AKM framework [Abowd et al., 1999, Card et al., 2013, Song et al., 2019] to decompose log-wages into components due to individual, i, and job, j effects, where jobs are defined by firm-occupation pairs $j \in \mathcal{J} = \mathcal{F} \times \mathcal{O}$.[3] Define the assignment function $J(i,t) = j$. This

[3]In our paper, we interpret the covariance as measuring worker-firm sorting in terms of pay premia in line with studies like Song et al. [2019], Card et al. [2013]. Recent research has shown that there is a less

specification is given in equation 1. Throughout, we also 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.[4]

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

This framework extends the worker-firm two-way fixed effects model in two dimensions. First, it allows firms to pay varying wage premia across occupations. For example, software developers can receive a higher premium than accountants within the same firm. Second, it allows firm-occupation-specific pay premia: software developers at Google can be paid a higher premium than software developers at other firms, even those with the same firm-pay premia.[5] The specification in equation 1 allows a simple decomposition of the observed variance of log wages as given in equation 2 below.

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

We wish to calculate the degree of worker-firm sorting using the decomposition given in equation 2. To do this, we focus on $\text{Cov}(\alpha_i, \lambda_j)$, which quantifies the degree of sorting between workers and jobs. Our aim is to capture the component of this variation that is due to sorting to firms and not occupations. Note that we can write $\lambda_j = \gamma_o + \psi_f + \Omega_{of}$, and therefore $\text{Cov}(\alpha_i, \lambda_j) = \text{Cov}(\alpha_i, \gamma_o) + \text{Cov}(\alpha_i, \psi_f) + \text{Cov}(\alpha_i, \Omega_{of})$. The first term of this expression captures sorting between workers and occupation fixed effects. We can decompose the second and third terms into that due to cross-occupation variation and that due to within-occupation, across-firm variation using the law of total covariance.

$$\text{Cov}(\alpha_i, \psi_f) = \mathbb{E}\left[\text{Cov}(\alpha_i, \psi_f|o)\right] + \text{Cov}\left(\mathbb{E}[\alpha_i|o], \mathbb{E}[\psi_f|o]\right) \tag{3}$$

$$\text{Cov}(\alpha_i, \Omega_{of}) = \mathbb{E}\left[\text{Cov}(\alpha_i, \Omega_{of}|o)\right] + \text{Cov}\left(\mathbb{E}[\alpha_i|o], \mathbb{E}[\Omega_{of}|o]\right) \tag{4}$$

straightforward relationship between worker fixed effects and firm productivity, both theoretically [Eeckhout and Kircher, 2011, Lopes de Melo, 2018], and empirically [Lochner and Schulz, 2024].

[4] 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.

[5] In particular, this framework is more flexible than a three-way fixed-effects specification with occupation, γ_o, and firm ψ_f fixed effects. To see this, note that we can write $\lambda_j = \gamma_o + \psi_f + \Omega_{of}$. Therefore, while the three-way fixed effects model captures the covariance between workers and occupations ($\text{Cov}(\alpha_i, \gamma_o)$), and workers and firms ($\text{Cov}(\alpha_i, \psi_f)$), it does not capture the covariance between workers and firm-occupation match effects ($\text{Cov}(\alpha_i, \Omega_{of})$. This missing component can be intuitively understood as the sorting of high-wage workers to high-wage firm-occupation pairs.

Intuitively $\text{Cov}(\mathbb{E}[\alpha_i|o], \mathbb{E}[\psi_f|o])$ is positive if occupations that attract high-wage workers are concentrated in high-wage firms, and $\text{Cov}(\mathbb{E}[\alpha_i|o], \mathbb{E}[\Omega_{of}|o])$ is positive if occupations with high-wage workers are concentrated at firms that reward those specific occupations more highly. Both of these terms are more likely to be non-zero as a consequence of the segregation of occupations across firms, and do not capture a notion of sorting between workers and firms. Therefore, we focus on $\mathbb{E}[\text{Cov}(\alpha_i, \psi_f|o)] + \mathbb{E}[\text{Cov}(\alpha_i, \Omega_{of}|o)]$ which it is easy to show is equal to $\mathbb{E}[\text{Cov}(\alpha_i, \lambda_j|o)]$, the within-occupation across-firm component of $\text{Cov}(\alpha_i, \lambda_j)$ in a law of total covariance decomposition. Intuitively, this object is positive if, after fixing occupation, high-pay workers systematically end up at high-pay firms or high-pay firm–occupation matches. It captures a notion of horizontal inequality, that is, inequality due to where you work, not what you do.

In a worker-firm model we estimate ψ_f^{AKM} and then quantify sorting from $\text{Cov}(\alpha_i^{AKM}, \psi_f^{AKM})$ in an occupation-blind manner. The difference between $S_{AKM} = \text{Cov}(\alpha_i^{AKM}, \psi_f^{AKM})$ and $S_{JFE} = \mathbb{E}[\text{Cov}(\alpha_i, \lambda_j|o)]$ is best described through an example. Suppose we have two private hospital firms, A and B, where hospital A hires mainly surgeons, whereas hospital B hires mainly nurses, and surgeons earn more than nurses. S_{AKM} would be large and positive due to the segregation of occupations across firms, even if both hospitals paid the same wage-premia. On the other hand, S_{JFE} would only be large and positive if sorting to hospitals occurred within occupation.

We next turn to the practical problem of recovering our object of interest, $S_{JFE} = \mathbb{E}[\text{Cov}(\alpha_i, \lambda_j|o)]$, from the data. Using the universe of matched employee-employer data, we can recover $\{\widehat{\alpha}_i, \widehat{\lambda}_j\}$ using equation 2. Note that fixed effects can only be recovered in a relative sense, and therefore, we can only estimate them for the largest connected set of jobs. S_{JFE} could be biased for two reasons. First, the fixed effects themselves could be biased, and S_{JFE} could inherit this bias. Second, as the fixed effects are only identified from movers across jobs, limited mobility bias will cause bias in quadratic transformations of the estimated fixed effects, such as S_{JFE}.

First, the estimates of the fixed effects are unbiased under the assumption of exogenous moves conditional on worker and job fixed effects. 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 main ways this assumption could be violated. 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 λ_j. Second, 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. 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. We probe these assumptions using event-studies and by focusing only on cross-firm moves following Card et al. [2013] in sub-section 3.1 and find that these concerns do not appear to be substantiated in our setting.

Second, turning to the limited mobility bias, we extend the leave-one-out variance estimation approach of Kline et al. [2020] deployed by Bonhomme et al. [2023], to our setting with conditional covariances. Intuitively, Kline et al. [2020]'s approach is to unbiasedly estimate the noise in the fixed-effects and use this to debias variance components using these noisy fixed-effects. We extend this simply by considering conditional covariance components [Azkarate-Askasua and Zerecero, 2024]. This approach has the disadvantage that we can now only estimate fixed effects in the largest, leave-one-out, connected set. This reduces the effective sample size and estimates results for a particular sub-sample of the data containing more well-connected workers and firms.

2 Data

We establish our empirical conclusions using administrative data from two of the largest labour markets in Europe: France and Germany. These datasets offer complementary strengths: the French data provides complete, uncensored wage information, but has to be constructed into a panel by stitching together individual cross-sections — a process with a roughly 95% success rate. The German data is available as a full-count panel, but the wage variables available are top-coded. Establishing our core results in both contexts, despite their different institutional settings and data strengths and weaknesses, underscores the robustness and generality of our findings. We compute results for the periods 2015-2019 for France, and 2017-2022 for Germany.[6]

Our empirical strategy leverages information about a worker's occupation; we thus describe how the occupation variable was collected separately for each country. In both cases, we check for the robustness of our results in two ways: (1) we use more aggregated occupational categories as there is less data error in the coding of more aggregate occupations, and (2) we estimate fixed effects based only on changes in occupation due to cross-firm moves, which are also less likely to be coded with error.

[6]For robustness, and to consider changes over time, we also look at past periods in both Germany and France in the online appendix.

2.1 French Sample

The main dataset underlying our French data is the "Base tous salariés" (BTS) data [Insee, 2024], a series of cross-sectional matched employer-employee datasets covering the universe of French workers except those in government employment. We follow Babet et al. [2022] in chaining these repeated cross-sections together into a quasi-panel tracking individuals over time, by matching their data in time t in the year t data to their information for time $t-1$ in the year $t+1$ data. This methodology allows for over 95% of individuals in each cross-section to be matched. One limitation of this approach is that those who are out of the labour force for more than one calendar year cannot be matched to their previous employment history and are instead given a new individual identifier.

To make our results comparable to the cross-setting study of AKM results in Bonhomme et al. [2023], our sample is restricted to full-time male workers aged 25-60 in metropolitan France, satisfying standard hours and earnings thresholds. Our decomposition is performed on log annual wages, residualized on a cubic in age and year fixed effects. The key variables that we use are real annual earnings, the firm identifier (SIREN), and the 4-digit occupational classification (PCS), which provides 430 distinct categories. A worker's occupation is collected from compulsory monthly employer surveys, where reported occupation titles are cross-checked against reported occupation codes using specialist INSEE software. In the 10% or so of cases where the codes do not agree, additional correction processes are used.[7]

2.2 German Sample

Our German data comes from the Employee Histories (BEH), which contains employment spells for all workers outside the civil service. Its primary limitation is the top-coding of wages at the social security ceiling, which we address by imputing censored wages using the established methodology of Card et al. [2013]. Key variables are real annual earnings, the establishment identifier[8], and the highly granular occupation codes based on the 5-digit KldB 2010 occupational classification (with 1286 categories in the full classification). Our sample selection mirrors the French data: full-time male workers aged 25-60 in West Germany. As with the French data, we use log annual wages residualized on an age cubic and year fixed effects.

The occupational variable is collected from filings by employers from social security records; since one's occupation is not important for this purpose, there is a perception that employers may not always update their workers' occupations in a timely manner. As a

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

[8]An establishment can be thought of as roughly a firm-industry group and is distinct from both firm and branch. See Card et al. [2013] for a discussion of this issue.

result, occupational moves within firms are likely to be under-reported.

2.3 Summary Statistics

Table 1 shows some summary statistics from the two most recent samples, that is, 2015-19 for France and 2017-22 for Germany. After implementing the restrictions described, our main sample covers 48.0 million observations consisting of 14.3m workers, 1.3m firms, and 4.5m jobs in France, and 54.6 million observations consisting of 14.2m workers, 1.4m firms, and 4.9m jobs in Germany.

One disadvantage of using the Kline et al. [2020] correction of second-order moments of estimated fixed effects is that it is only feasible for the leave-one-out connect set (LOO) for the worker-jobs model. We present summary statistics for the connected sub-sample, for which identification of the model is feasible, in column 2 and for the leave-one-out connected subsample, for which the KSS correction is feasible in column 3.

Once we consider the leave-one-out connected set for the worker-jobs model, we have significantly fewer observations (around 66% of the overall total for the French data and 60% for the German data). 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 in France (56% of workers, 29% of establishments, and 19% of jobs in Germany). On the other hand, we find that the mean and variance of earnings is similar across the samples. The mean in the leave-one-out sample is about 0.04 log points higher in terms of annualised wages in France and Germany, and the residualised log-wage variance is somewhat smaller, by 0.03 and 0.01 respectively.

We also provide summary statistics on the number of moves we observe in the data. 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 that are between occupations, and the number of moves that are between both firms and occupations. In the full data, we find that 14.5% of observations involve moves in France and 10.6% of observations involve moves in Germany. Of these moves, 51.6% involve moves between firms, 79.0% involve moves between occupations, and so 30.5% involve changes in both firms and occupations (84.2%, 65.1% and 49.4% for Germany).

We find that the ratio of moves to observations in the LOO connected set is comparable to that in the full data. For the LOO connected set, 14.1% of all observations involve moves, of which 57.1% are moves between firms, 74.5% are moves between occupations and 31.8% are moves between both firms and occupations in France (10.9%, 84.1%, 61.7% and 45.8% in Germany). Thus, despite only containing 2/3 of the full sample, the leave-one-out connected sample nevertheless resembles the full sample in terms of the composition of moves across firms and occupations.

Table 1: Summary statistics for the French and German samples

	Full Sample	Connected Sample	Leave-one-out Connected Sample
France (2015-2019)			
Total observations (m)	48.01	35.93	31.53
Total workers (m)	14.25	9.06	8.09
Total firms (m)	1.25	0.57	0.27
Total jobs (m)	4.48	2.04	0.82
Mean log annual wage	10.39	10.42	10.43
Var log annual wage	0.25	0.23	0.23
Mean log hourly wage	2.96	2.97	2.98
Var log hourly wage	0.21	0.20	0.19
Var residualised log hourly wage	0.24	0.22	0.21
N moves (m)	6.97	5.69	4.43
N firm moves (m)	3.59	3.25	2.53
N occ moves (m)	5.50	4.38	3.30
N firm + occ moves (m)	2.13	1.94	1.41
Germany (2017-2022)			
Total observations (m)	54.56	40.16	32.96
Total workers (m)	14.22	9.44	7.95
Total estabs (m)	1.35	0.78	0.39
Total jobs (m)	4.91	2.30	0.92
Mean log daily wage	4.78	4.80	4.82
Var log daily wage	0.27	0.26	0.26
Var residualised log daily wage	0.25	0.24	0.24
N moves (m)	5.79	5.00	3.58
N estab moves (m)	4.88	4.27	3.01
N occ moves (m)	3.77	3.29	2.21
N estab + occ moves (m)	2.86	2.56	1.64

Notes: This table presents summary statistics for the two most recent periods studied in both Germany and France. Statistics on the number of observations and number of moves are presented in millions. Column 1 presents statistics for the full sample after the restrictions described in section 2, while columns 2 and 3 describe statistics for the largest connected set and the largest leave-one-out connected set, respectively.

2.4 Segregation of Occupations Across Firms

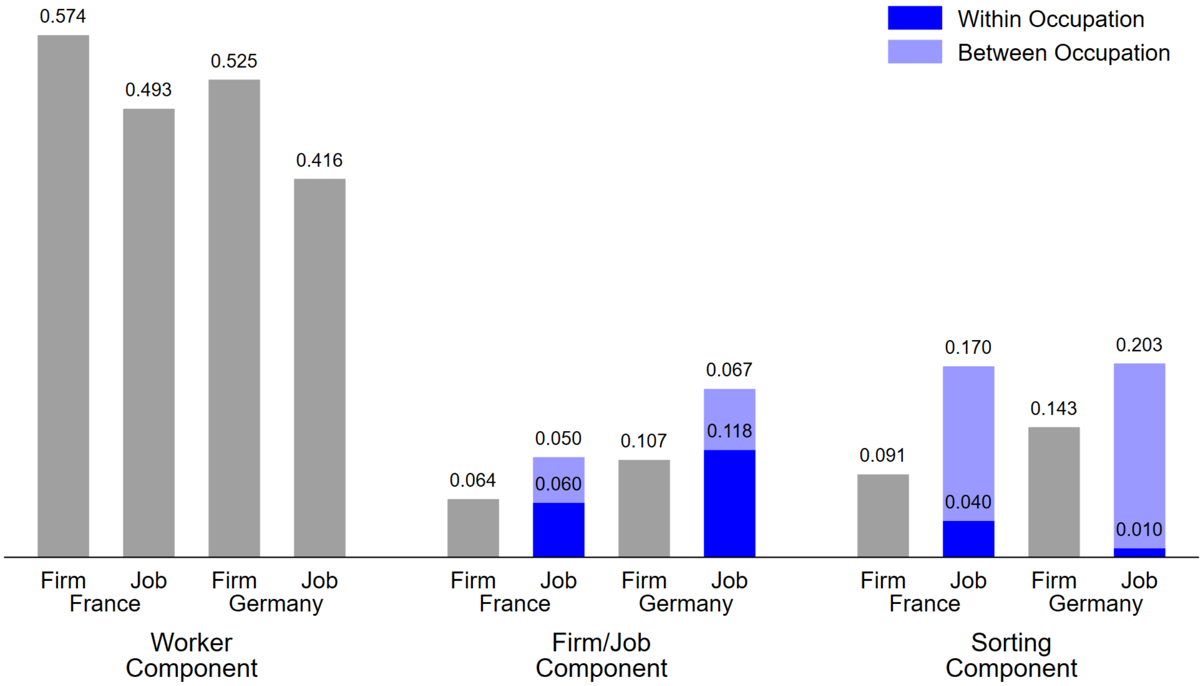
It is important to account for occupations and firms jointly because occupations are segregated across firms. To quantify the extent of this segregation, we turn to information theory and compute the Thiel-H index across firms [Theil and Finizza, 1971, Theil, 1972]. The Thiel-H index is a measure of segregation between multiple groups and can be interpreted as how much uncertainty about a worker's occupation would be resolved if one knew which firm they worked at. If occupations were perfectly segregated across firms, then knowing a worker's firm would be equivalent to knowing their occupation. If, on the other hand, occupations in each firm reflected the overall distribution of occupations in the economy, knowing the worker's firm would not improve one's guesses about their occupation. We present the index normalised with respect to the overall amount of uncertainty in workers' occupations, so that the index lies between 0 and 1. To compute the index, we use firms with more than 10 workers, so as to ensure that the findings are not driven solely by very small firms.

We find that occupations are highly segregated across firms in both Germany and France. Knowing what firm a worker works at removes about two thirds of the uncertainty about their occupation in both Germany and France. While reducing the fineness of occupational classification lowers the measured segregation, the effect is not substantial; for example, at the 3-digit level, knowing one's firm still removes about 63.2% of uncertainty in Germany and 62.7% in France. The online appendix shows details of these results, and that segregation has increased somewhat over time in both settings.

3 Decomposition results

Figure 1 shows the main decomposition results for both the French and German data side-by-side. It shows the results from decomposing wages using the fixed effects estimated from equation 1, correcting for limited mobility bias and using the law of total variance and covariance to separate between-occupation and within-occupation across-firm components. Figure 1 also shows the results from employing the same procedure on the standard AKM worker-firm model for comparison.

Figure 1: Decomposition Results



Notes: This figure shows our main decomposition results. Each bar presents the share of total residualised log-wage variance of the described component for the specified model. For the variance of job-fixed effects and the covariance of worker- and job-fixed effects, we show its decomposition into between-occupation and within-occupation components as described in the text.

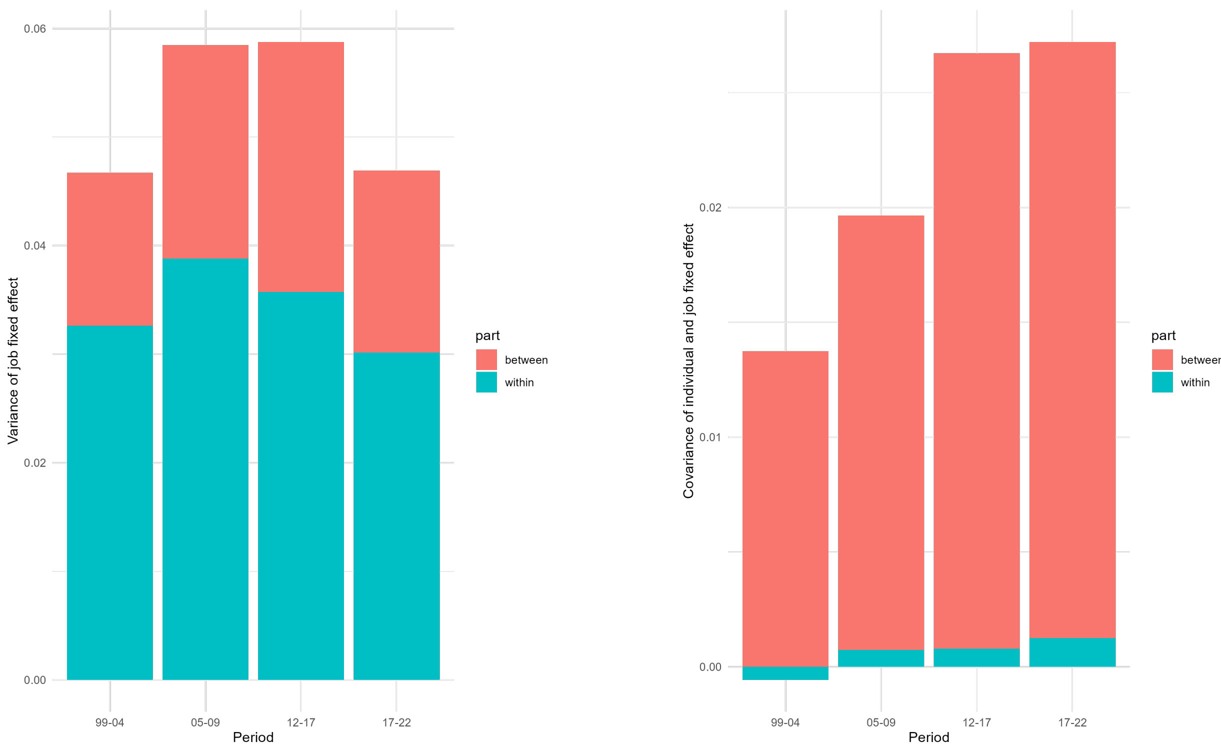
We find that the sorting of workers to firms within occupations accounts for a small proportion of log-wage variance (4% in the French sample, and 1% in the German sample), significantly less than that recovered by a worker-firm AKM model. On the other hand, sorting across occupations accounts for a relatively large proportion of log-wage variance (17% in the French sample and 20% in the German sample). Furthermore, worker-occupation sorting may underlie much of what was previously attributed to worker-firm sorting in decompositions based on the standard AKM model. The sorting of workers to firms within occupations is only 44.0% as large as the estimated sorting of workers to firms in the AKM model in France and only 7.0% as large in Germany.

Figure 1 also shows that firm heterogeneity explains a similar proportion of log wage variance in the worker-job and worker-firm models in both France and Germany (in France around 6%, whereas in Germany around 11%). In both countries, we also find a significant role for occupation heterogeneity (5% in the French sample, and 6.7% in the German sample). Our results thus support the findings in prior work that there is substantial wage premia between firms even within finely defined occupation groups. However, there is little sorting of workers to firms within occupations, even given this variance.

Finally, in both France and Germany, we find that worker-heterogeneity explains significantly less log wage variance in the worker-job as opposed to the worker-firm model, reducing from 57% to 49% in France and from 53% to 42% in Germany.

Finally, the literature has recently been interested in how the role of firms in explaining wage inequality has changed over time (e.g. Song et al. [2019]). In the French setting, due to data restrictions, we cannot go far back in time. However, using the German data, we can perform our decomposition every five years since 1999; we report the time-varying within-between decomposition in Figure 2b.[9] We find that the degree of sorting between workers and jobs has increased over time. While we confirm that sorting between workers and firms within occupations has increased over this period (as previously documented in Card et al. [2013]), we also find that the rise is mainly attributable to greater sorting between workers and occupations.

Figure 2: Changes in variance of job fixed effects and sorting from 1999-2022 in Germany



(a) Variance of job fixed effects

(b) Covariance of individual and job fixed effects

Notes: These figures plot the variance of job fixed effects (in panel 2a) and the covariance of individual and job fixed effects (in panel 2b. We decompose these terms into between-occupation (orange) and within-occupation (blue) terms as described in the text. We plot these statistics for four time periods in Germany, 1999-2004, 2005-2009, 2012-2017, 2017-2022. The years 2010 and 2011 are excluded because of an occupational change which led to abnormally large numbers of occupational moves.

[9] In 2010-11, a new occupational classification was introduced in Germany, inducing an abnormally large number of apparent occupational moves. We exclude those years from the analysis because occupation changes in these years will not always reflect actual occupational moves.

3.1 Diagnostics and robustness checks

In this section, we discuss diagnostic tests of the econometric assumptions underlying our interpretations in section 3 and probe the robustness of this analysis to various specification changes. In addition to the analysis described in this sub-section, we find that the qualitative takeaways of our main decomposition result are also robust to using log hourly wages instead of annual wages, as well as using a sample of only women. These results are available on request.

3.1.1 Exogeneity of moves

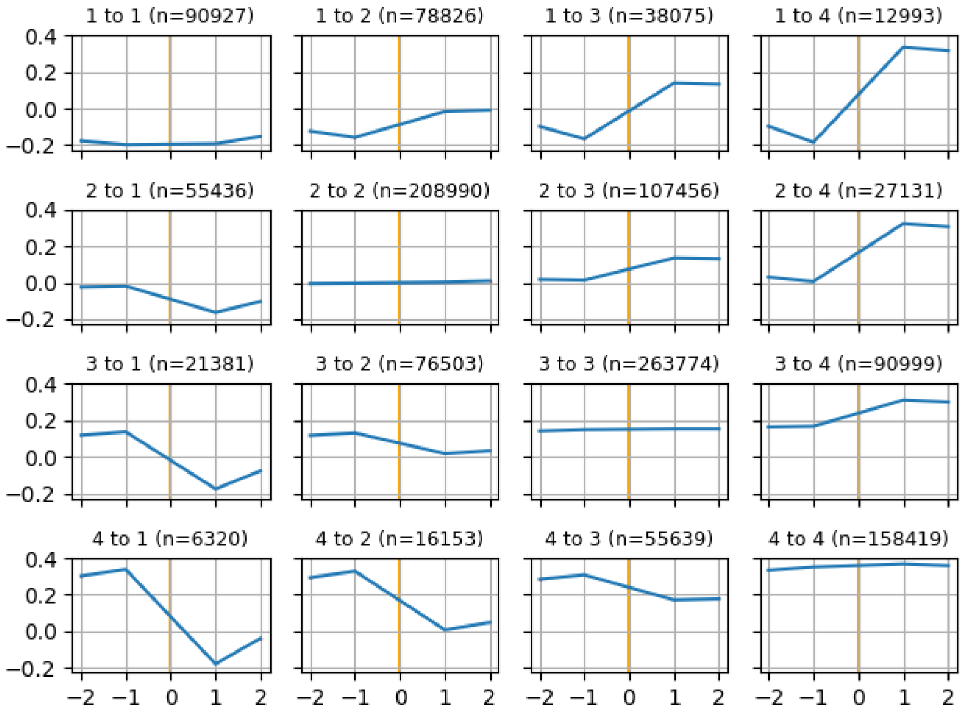
As described in section 1, to credibly interpret the job fixed effects as pay premia, we require exogenous moves of workers to jobs conditional on the estimated fixed effects. We follow the logic of Card et al. [2013] by considering wage changes around job changes. If the econometric model specified 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 not to lead to a symmetric effect on earnings. Card et al. [2013] produces 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 implement these event-study "tests" by categorising the jobs that individuals move into or out of into four earnings quartiles using the leave-out job-specific mean fixed effect. Figure 3 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 categorising jobs using average job wages in the online appendix, which shows similar results. Symmetric 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].

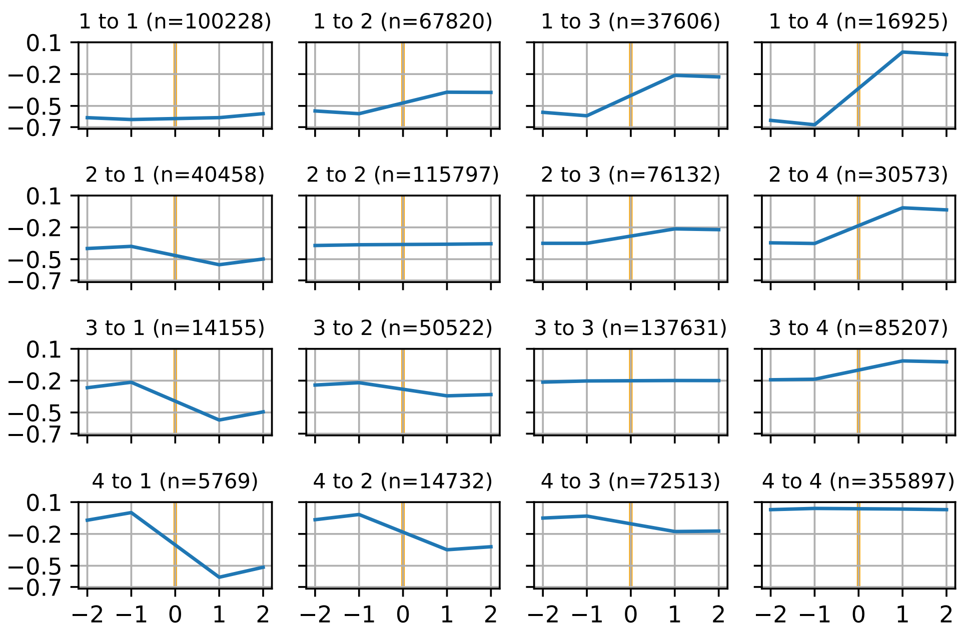
Finally, one might be worried that the averages presented in the event study might obfuscate different degrees of exogeneity between different types of moves; e.g. moves across firms may be conditionally exogenous as is accepted in the AKM literature, while moves across occupations within a firm may not be, particularly if firms have more private information about workers which allow them to achieve significant match effects. We therefore plot a version of the event study using only moves within firms across occupations in online appendix. Even for these moves, we still observe the symmetric wage gains and losses implied by the two-way fixed effect model when workers move across occupations in different fixed effect quartiles.

Figure 3: Event study around job moves, clustering by leave-out job mean fixed effect

(a) France



(b) Germany



Notes: These figures show the impact on average wages around the event of job movement. Each cell shows the average wage change associated with a movement event from one 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. Panel (a) shows the results for France, and panel (b) shows the results for Germany.

3.1.2 Linearity of worker and job effects

A possible criticism of the approach we take 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 again follow the approach in Card et al. [2013] and show in the online appendix that the mean of the residuals of the two-way FE regression are near zero on average, and by job and worker cells. We find that the specification performs well in this test.

3.1.3 Robustness to coarser occupation categories

Another concern is that occupations could be measured with error in our data. First, if we use very fine-grained occupational codes, it could be difficult for firms to consistently fill in workers' occupations accurately. Apparent segregation across firms could reflect differences in reporting standards and practices across firms and not actual differences in the work done. To check for this, we reproduce our main results using coarser occupation definitions; whereas at the 4- or 5-digit level mistakes in occupation classification may be made, this seems more improbable the more aggregated occupation groups become.

Naturally, as the occupation classification becomes coarser, across-occupation differences will mechanically explain less variation, even in the absence of measurement error. However, we find that even at the one-digit occupation level in the French setting, 63% of the wage inequality that can be attributed to sorting between workers and jobs is due to sorting across occupations, and 14% of the wage inequality that can be attributed to job-heterogeneity is due to across-occupation differences. Details of results and decompositions for one-, two-, three-, and four-digit occupation classifications are given in the online appendix.

3.2 Discussion

Our results offer a different perspective on log wage inequality than most of the literature. The initial development of the AKM model drove a lot of research into firm wage premia Mortensen [2003], and subsequent landmark papers like Card et al. [2013] and Song et al. [2019] have reinforced the importance of sorting in contributing to wage inequality. Our results qualify these findings by showing that sorting across occupations is quantitatively more important than sorting across firms within occupations in explaining wage inequality. Thus, our results suggest that mechanisms driving occupational choice are particularly important in driving wage inequality. Our results are consistent with classic Roy-type selection on comparative advantage and task returns, e.g. [Roy, 1951, Acemoglu and Autor, 2011, Yamaguchi, 2012], suggesting that some of this inequality might be driven by efficient matching. However, to the extent that family background might affect both worker ability (e.g. through better education) as well as occupational access, our results are also consistent with papers that suggest that family background is important in regulating access to desirable occupations [Aina and Nicoletti, 2018, Bell et al., 2019, Bamieh and Cintolesi, 2021, Lo Bello and Morchio, 2022, Almgren et al., 2025].

Second, our results have confirmed the existence of large variation in firm wage-premia even within occupations, yet we find little sorting of workers across firms within occupations. This raises a puzzle: if some firms pay more within the same occupation, why don't higher-skill workers systematically match to those higher-paying firms?

One explanation is that this finding is not fully at odds with the theoretical literature; after all, Eeckhout and Kircher [2011] and Lopes de Melo [2018] point out that theoretical models of sorting on productivity do not imply a strictly increasing relationship between worker and firm pay premia. Furthermore, part of the story could be that wage differentials across firms might represent compensating differentials; firms that offer higher wages may offer fewer non-wage amenities or require work in less desirable locations.[10] Another possible explanation of this phenomenon is our setting — France and Germany have relatively rigid labour markets. Without the ability to easily let unproductive workers go, firms may be unable to break unproductive matches and form new, productive ones. To this end, it would be interesting to see if the facts documented in this paper can be replicated in countries with more flexible labour markets.[11]

4 Conclusion

In this paper, we argue that the degree to which high-wage workers sort to high-wage firms has been overestimated in AKM-style decompositions. This overestimation is due to high-wage occupations clustering in high-wage firms. Not explicitly accounting for a worker's occupation, therefore, leads to sorting to high-wage occupations being mistaken for sorting to high-wage firms.

We extend the standard AKM model by estimating worker-job two-way fixed effects instead of worker-firm two-way fixed effects, where jobs are occupation-firm pairs. We show using event studies around job changes that wages experience step-changes consistent with the fixed effects model when they move from high-wage jobs to low-wage jobs. We account for limited-mobility bias using the leave-one-out variance estimator due to Kline et al. [2020],

[10] The empirical evidence on this question is mixed: Humlum et al. [2025] finds a negative correlation between pay and non-wage amenities in Denmark, while Sockin [2022] finds that higher wage firms also offer better wage amenities.

[11] To our knowledge, the data underlying AKM studies based in the US does not contain information on occupations.

and demonstrate robustness of our core results to using coarser occupation codes, considering different time periods, and considering different data definitions.

We show that quantitatively, sorting of workers to occupations accounts for far more of the total log wage variance than sorting of workers to firms within occupations. Estimates of worker-firm sorting from standard AKM models are substantially higher than estimates of sorting of workers to firms within occupations in our model, suggesting that much of what was previously considered worker-firm sorting may have been worker-occupation sorting instead. Second, we show that even after accounting for occupations, there is considerable variation of firm wage premia within occupations. Thus, our results suggest fairly low sorting of workers to higher-paying firms despite the existence of relatively large firm wage premia dispersion.

References

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

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.

Carmen Aina and Cheti Nicoletti. The intergenerational transmission of liberal professions. *Labour Economics*, 51:108–120, 2018.

Mattias Almgren, John Kramer, and Jósef Sigurdsson. It runs in the family: Occupational choice and the allocation of talent. Technical report, CESifo Working Paper, 2025.

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.

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.

Omar Bamieh and Andrea Cintolesi. Intergenerational transmission in regulated professions and the role of familism. *Journal of Economic Behavior & Organization*, 192:857–879, 2021.

Alex Bell, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics*, 134(2):647–713, 2019.

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.

Jan Eeckhout and Philipp Kircher. Identifying sorting—in theory. *The Review of Economic Studies*, 78(3):872–906, 2011.

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.

James J Heckman and Guilherme Sedlacek. Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market. *Journal of political Economy*, 93(6):1077–1125, 1985.

Anders Humlum, Mette Rasmussen, and Evan K Rose. Firm premia and match effects in pay vs. amenities. Technical report, National Bureau of Economic Research, 2025.

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

Michael P Keane and Kenneth I Wolpin. The career decisions of young men. *Journal of political Economy*, 105(3):473–522, 1997.

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 and Fabien Postel-Vinay. Multidimensional sorting under random search. *Journal of political Economy*, 131(12):3497–3539, 2023.

Salvatore Lo Bello and Iacopo Morchio. Like father, like son: Occupational choice, intergenerational persistence and misallocation. *Quantitative Economics*, 13(2):629–679, 2022.

Benjamin Lochner and Bastian Schulz. Firm productivity, wages, and sorting. *Journal of Labor Economics*, 42(1):85–119, 2024.

Rafael Lopes de Melo. Firm wage differentials and labor market sorting: Reconciling theory and evidence. *Journal of Political Economy*, 126(1):313–346, 2018.

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

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

Andrew Donald Roy. Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146, 1951.

Michael Sattinger. Assignment models of the distribution of earnings. *Journal of economic literature*, 31(2):831–880, 1993.

Robert Shimer. The assignment of workers to jobs in an economy with coordination frictions. *Journal of political Economy*, 113(5):996–1025, 2005.

Robert Shimer and Lones Smith. Assortative matching and search. *Econometrica*, 68(2): 343–369, 2000.

Jason Sockin. Show me the amenity: Are higher-paying firms better all around? 2022.

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.

Henri Theil. Statistical decomposition analysis: With applications in the social and administrative sciences. *(No Title)*, 1972.

Henri Theil and Anthony J Finizza. A note on the measurement of racial integration of schools by means of informational concepts. 1971.

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.

Shintaro Yamaguchi. Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30(1):1–53, 2012.