Journées de méthodologie statistique de l'Insee 2022

Pour qui sont les bons salaires ? Une estimation débiaisée du rôle des appariements dans la dynamique des inégalités de salaire

Damien BABET (*), Olivier GODECHOT (**), Marco Guido PALLADINO (***)

(*) Insee, Direction des études et synthèses économiques
(**) CNRS, Sciences-Po
(***) Sciences-Po

damien.babet@insee.fr

Mots-clés. : Données de panel, effets fixes, split sampling, surapprentissage, formes quadratiques.

Domaines. Économétrie.

Résumé

En Allemagne et aux États-Unis notamment, la croissance des inégalités de salaires se manifeste surtout entre entreprises plutôt qu'au sein des entreprises. Cela s'explique par les différences entre entreprises, mais aussi par l'effet de sorting ou d'appariement : certaines entreprises payent mieux, et ces employeurs ont de plus en plus tendance à embaucher des salariés qui gagnent plus (Card, Heining, Kline 2013, Song et al. 2019). Cette augmentation de l'effet d'appariement se manifeste-t-elle également en France ?

La mesure du *sorting* se fonde sur un modèle de panel à doubles effets fixes de type « AKM » (Abowd, Kramarz et Margolis, 1999) où le logarithme du salaire d'un salarié i est la somme de l'effet entreprise $\psi_{j(i,t)}$ de l'employeur j de i et d'un effet individuel θ_i, capturant l'hétérogénéité inobservée, ainsi que de l'effet de covariables x qui ne sont pas constantes dans le temps.

$$y_{it} = \beta x_{it} + \theta_i + \psi_{j(i,t)} + u_{it}$$

Le sorting est alors la covariance entre les effets entreprises et les effets individus $cov(\theta_i, \psi_j)$. Or l'identification de ces effets fixes s'appuie sur les salariés mobiles et la variation de leur salaire lorsqu'ils passent d'un employeur à l'autre. Il existe alors un biais « de mobilité limitée », un phénomène de surapprentissage déjà connu, du fait que la connectivité du réseau des entreprises, et l'estimation des effets fixes relatifs, peut reposer sur un faible nombre de salariés mobiles : la variance des effets fixes est biaisée à la hausse et leur corrélation à la baisse. De manière générale, ce sont toutes les formes quadratiques des estimateurs qui sont biaisées.

Deux principales solutions sont aujourd'hui proposées dans la littérature : le *clustering* des entreprises ou des individus pour réduire la dimension de l'hétérogénéité des effets fixes (Bonhomme, Lamadon, Manresa 2017), qui demande des hypothèses additionnelles, et l'estimation de l'erreur d'estimation de chaque effet fixe par *leave-out* (Kline, Saggio Sølvsten, 2020), qui reste complexe et coûteuse en temps de calcul.

Nous introduisons deux innovations. D'une part, nous estimons pour la première fois l'évolution du sorting sur des données françaises, grâce à la constitution d'un pseudo-panel en chaînant les DADS exhaustives (Godechot et al. 2020). D'autre part, nous utilisons une méthode de *split-sampling* dont nous montrons qu'elle est valide sous les mêmes hypothèses que la méthode de *leave-out*, mais plus simple et rapide. On peut en effet écrire l'estimateur *split-sampling* d'une forme quadratique des paramètres comme $\hat{\omega}^{SP} = \hat{\alpha}_0' A \hat{\alpha}_1$ avec $\hat{\alpha}_s = (\hat{\beta}, \hat{\theta}, \hat{\psi})_s$, $s = 1, 0$ indiquant la moitiée d'échantillon utilisée et A étant une matrice symétrique appropriée. Avec ϵ l'erreur d'estimation des paramètres, $\hat{\alpha}_s = \alpha + \epsilon_s$ on a alors en espérance un biais égal à la forme quadratique des erreurs entre chaque moitié de l'échantillon, qui est nulle sous des conditions raisonnables :

$$\begin{aligned}\mathbf{E}[\hat{\omega}^{SP}] - \omega &= trace(A\mathbf{E}[\epsilon_0 \epsilon_1']) \\ &= \mathbf{E}[\epsilon_0' A \epsilon_1]\end{aligned}$$

Cette correction augmente très sensiblement la mesure de l'effet de *sorting*. Estimé ainsi, l'appariement explique 11,6% de la variance des log-salaires en France sur la période 2002-2006, et 13,5% sur la période 2012-2016. Toutes les estimations, corrigées ou non, s'accordent pour décrire une augmentation du poids du *sorting*.

Abstract

Increase in wage inequalities in the USA and Germany are explained in part by a rise in the sorting of high-wage workers to high premium firms (Card, Heining, Kline 2013, Song et al. 2019). Such a result comes from a decomposition of the log-wage variance into variance of firms and workers fixed-effects obtained, and their covariance, obtained from an estimation of the AKM model (Abowd, Kramarz et Margolis, 1999). Thanks to a newly built quasi-exhaustive panel of French workers, we extend these results to France, showing that, despite stable wage inequalities between 2002 and 2016, sorting also increased. We further add to the literature by proving the efficacy of a split sampling method to correct for the limited mobility bias in sampling, and adapt variance decomposition methods to covariance decomposition to explore the causes of the rise in sorting. We show that this rise is robust, and linked both to firm demographics and to changes in the distribution of occupations among firms.

Introduction

Wage inequality is a driving force of economic inequalities. Its rise for several decades in most rich countries is well documented[1]. Firms play a central role in this rise : in Germany (Card et al. (2013)) and in the USA (Song et al. (2019)) rising inequality comes in large part from between-firm inequalities. Both papers use Abowd et al. (1999) (hereafter AKM) model of log-wages with additive workers and firm fixed effects to decompose log-wage variance into three main components : variance in individual workers heterogeneity, variance in firm premium, and

1. Tomaskovic-Devey et al. (2020) for a recent international comparison

covariance between the two, or *sorting*. In both cases, sorting explains a large share of the rise in wage inequalities. High-wage workers tend to work for high-wage firms, and increasingly so.

This paper provides a similar decomposition of log-wage variance and its evolution for France between 2002 and 2016, using a new dataset of almost exhaustive matched employers/employees wages for France, described first in Godechot et al. (2020). We show that sorting increased as in other countries. We further add to the literature in two main ways : we introduce and prove the efficacy of a split sampling method to correct for the limited mobility bias, and compare it to the Bonhomme et al. (2019) model estimates. We then introduce covariance decomposition to investigate the rise in sorting, and find it explained in great part by a change in the distribution of occupations between firms.

France is an interesting touchstone because it is often considered an exception : wage inequality there has been stable or decreasing on the last decades. Bozio et al. (2020) show that since the 1970's the redistributive effects of payroll taxation has regularly increased. Considered before tax, labour costs inequalities have increased in France at a comparable rate as in other countries. Regarding the rise in sorting also France is no exception, suggesting the same mechanisms raising inequalities elsewhere are also at work in France, only to be balanced by increasing redistribution. With a different focus, Godechot et al. (2020) measure work segregation based on wages in eleven countries and find it growing in most, including France.

The measure of sorting through AKM models has however a well-known "limited mobility" bias described in Abowd et al. (2004), Andrews et al. (2008), and Bonhomme et al. (2020). This bias stems from the limited number of useful observations for each individual firm and worker parameters. Individual parameters estimations remain consistent, but the variance of the error term is underestimated. This bias on the quadratic forms of estimated parameters exists in all linear regression but is made worse by the low number of useful observations for some parameters. When a firm's fixed effect is overestimated, its employee's fixed effects will tend to be underestimated. The high estimation errors are correlated between parameters in complex ways dependent on the network structure (Jochmans and Weidner, 2019). This directly affects the estimation of variance components, and it directly biases negatively the estimation of the correlation between firms and workers effects. Both Card et al. (2013) and Song et al. (2019) acknowledge this important bias in the measure of sorting, but expect it to be stable enough in time that it does not impact their dynamic results.

Several correction strategies are available. Andrews et al. directly correct estimates with a bias correction factor derived from an estimate of the error term variance, with an hypothesis of homoscedasticity that is unrealistic because of the networked nature of the estimation error. Borovičková and Shimer (2017) model heterogeneity as random effects rather than fixed effects and find a much higher sorting that previous estimates, but while fixed effects models allow for further study of the distribution of this heterogenity, it is much more difficult with random effects models. Bonhomme et al. (2019) cluster firms based on the distance between their wage distributions, then estimate a wage model where workers' effects are treated as random effects. The clustering allow for richer models, including interaction and dynamic terms, at the price of the additional hypothesis that clusters are correctly identified. Kline et al. leave-one-out strategy amounts to the bias correction factor method of Andrews et al. compatible with heteroscedasticity, but is complex and computationally costly on large datasets. We use a much simpler split-sampling strategy by applying only one split to our data, rather than a leave-one-out method with as many splits as observations. Split-sampling has been used in similar settings by Chanut (2018), Drenik et al. (2020), Goldschmidt and Schmieder (2017), Gerard et al. (2018), and Schoefer and Ziv (2021). We add to these works by generalizing the idea and providing a proof. We also implement Bonhomme et al. (2019) cluster method (without random effects) and find the results coherent with split sampling. We also implement firm clustering with random effects to check for interaction and dynamic mobility patterns, and find these patterns to be of limited impact.

Once the rise in sorting is robustly estimated, it remains to be explained. We regress the AKM-estimated firms and workers fixed effects on explanatory variables at the firm level (firm size, value added, share of women in the workforce, age distribution, occupation distribution) and use the estimated parameters to decompose the covariance in out first and last periods. We did not find precedent for this extension of variance decomposition. Our explanatory models of fixed effects, however, are limited by the information we have and likely victims of an omitted variable bias. We exploit the panel structure and add a new firm fixed effect (a meta fixed effect) in the regressions to account for firm selection in omitted variables correlated to our explanatory variables. We thus obtain a covariance decomposition that arguably reflects the impact on sorting of the shifts in firm sizes, shares of women, occupational distribution, etc. We find a central but complex role for occupations, as well as differentiated growth dynamics for firms depending on their role in sorting.

1 Data

1.1 Building an exhaustive pseudo-panel

We use DADS data, exhaustive yearly files built from tax returns files by firms on their payrolled employees. The data is pseudonymous with an identifying code that changes every year, allowing for cross-section use of the file, but not for long panel use. Panel analysis on French matched employers / employees wage data are traditionally done on the "DADS panel" or "all wage-earners panel", which we dub here the "narrow panel". This panel is build on a sample of 1/24 before 2002 and 1/12 after, sampling the same individuals as a permanent demographic panel, allowing matching. The sampling also allows for additional data quality control and correction work that would be much heavier on the exhaustive data. This narrow panel was the basis for the original AKM.

Since then, though, AKM models have been better estimated in countries where researchers have had access to exhaustive panel data : USA, Germany, Sweden, Austria, Italy, Norway, Denmark, etc. and for good reason. As in any estimation, the reduction and the sample size raises uncertainty, so that, for instance, firms in the narrow panel need to be roughly 12 times bigger to have their fixed effect estimated with the same precision as in the exhaustive data. Moreover, AKM estimation relies on mobile workers moving between firms, and is only possible conditional on the group of firms interconnected through such workers. Sampling drastically reduces the proportion of firms belonging to the main connected component, and further reinforces the selection effect of bigger and more connected firms in the estimation sample. All these problems translate into an even larger "limited mobility bias" for variance and sorting estimates.

Fortunately, each yearly DADS file is also a short panel, with most job variables given both for the year t as well as for the year $t-1$. A direct use of this data as short, two years panel data is possible, but this overlap also allows for matching between yearly files, based on common information (establishment ID, gender, number of hours, job duration in days, start and end dates of the job, municipality of work and residence, earnings and age) between year t of yearfile y–1 and year t–1 of yearfile y. Between 2002 and 2016, matching gives a single match with 98% of the individuals. Matching misses include specific situations where all matching variables are identical for several individuals (such as higher education institutions for civil servants for instance), and rares instances of individual data modifications between one yearly file and the $t-1$ values of the following year.

By construction, even if matching were perfect, employment spell before and after a total career interruption of more than a year cannot be connected to the same individual. Employees who are not matched either due to career interruption or to matching misses still are in the panel, but they appear under multiple ident numbers. We dub the resulting almost exhaustive pseudo panel the "wide panel".

This matching procedure is of no use before 2002. Up to 2001, the various jobs of a single individual were not linked, in the files, by a corresponding individual ident number. It is not possible to follow one workers through different employers, even in the course of one year. Matching is possible, but only for workers who kept the same job during the two years to be matched. Since estimation of AKM relies on the different wages a given individual can earn when working for different employers, it can not be done before 2002.

We nevertheless computed long term series on sorting from 1976 to 2016, relying on the small panel, where information on wages and careers is available since 1976, with some missing years (1981, 1983, 1990) and varying data quality.

We mobilize other data sources. Exhaustive firm financial data from administrative sources (FICUS/FARE files) matched to our wage files provide value-added per worker and total workforce, irrespective of the sample restriction we use. These variables are exploited in the covariance decomposition to explain the rise in sorting and prove valuable.

1.2 Sample restrictions

Following similar works, we exclude public workers, both for comparability and data quality reasons. We also restrict to ordinary jobs, excluding subsidized contracts, interns and apprenticeship. We include both men and women.

We divide the data in three five-years periods : 2002-2006, 2007-2011 and 2012-2016. Each observation consists in a worker / firm / year triplet, where each individual worker is associated with the firm from which she earned the most during the year (or, when equal, for which she worked the most). For simplicity, we sometimes call such observations a "wage", or a "job". Each worker can appear up to 5 times during each of the sample periods.

We have information about the number of hours worked, which is rare in this kind of data. Without it, it is common practice in the literature to set a minimal wage for inclusion and exclude women to reduce the risk of misidentifying part-time workers. We can avoid these exclusions. Consequently, our target variable of earnings is the log hourly wage. We restrain the sample to people employed for the full year, so as to limit the impact of annualized payments, and because for people with spells of unemployment or inactivity, the total annual number of hours worked (as well as the duration in days) is not entirely reliable and computation of hourly wage lacks precision [2]. We further exclude jobs with hourly wage inferior to 80% of the legal minimum hourly wage for the corresponding year, or above 1000 times the minimum hourly wage, and observations with missing values for sex, age and employer. All restrictions are done after matching, when the wide panel is already constructed. They select specific observations but not individuals, who remain in the wide panel as long as they have been working during the year (or received unemployment benefits).

In our historical series computed on the small panel, work hours are not reliable before 1996, but job duration in days and an indicator variable for part-time jobs exist since 1976. In our long-term series we also compute sorting on the daily wage of full-time workers. The level is lower than hourly wage small panel sorting, but the trends are almost parallel for the post-1996 period where both are known. Other changes in variables in the 40 years period also preclude an exact reconstruction of the selection we choose on the wide panel, notably because the distinction between public and private sector is not consistent.

2. We found similar results when widening the selection to all individual whose main job during the year lasts more than 90 days. In this larger sample moreover, the connectivity is better and the limited mobility bias is reduced

2 Methodology

2.1 AKM model

We follow AKM with an additive model of log-wages :

$$y_{it} = \beta x_{it} + \theta_i + \psi_{j(i,t)} + u_{it} \tag{1}$$

Here y_{it} is the logarithm of the hourly wage of worker $i = 1, 2, ..., N$ during year $t = 1, ..., T$, demeaned by the average log-hourly wage for all workers during year t so that $\bar{y}_t = 0$ [3]. Time-varying covariates x_{it} are limited to age and age squared. θ_i is the fixed effect of individual worker i, and ψ_j is the fixed effect of individual firm $j = 1, 2..., J$, both supposed constant in time for the duration of the panel, firm $j(i,t)$ being the employer of worker i during year t. u_{it} is the idiosyncratic error term. We further note $\mathbf{F} = (\mathbf{1}_{j=j(i,t)})$ the $N^* \times J$ matrix of the bipartite graph of workers/firms connections through time, with $N^* = NT$.

This model rests on two notable hypothesis :

— No interaction effect between firm type and worker type : the fixed effects are additive (in the log-wage). We suppose the firm specific wage premium will be the same for all workers, men and women, young and old, skilled or not.

— Exogenous mobility : the residual term u_{it} has null expectation conditional on the variables x_{it}, i, t and j, as is classical, but also conditional on the matrix $\mathbf{F}$. This means in particular that wages before or after a job change are on average the same as if there had been no job change.

Although both hypothesis can appear unrealistic and have been subjected to scrutiny, they seem to provide reasonable approximations. Bonhomme et al. (2019) build a model that allows for interaction and endogenous mobility and find only slight departures to the additive linear model. We replicate their model on our data and reach the same conclusion.

More recently and more generally, De Chaisemartin and d'Haultfoeuille (2020) show in a different context that heterogeneity in treatment effect can be a serious obstacle to correct identification of the average treatment effect. Their analysis does apply to the AKM setting, if one considers for instance each firm fixed effect as a treatment effect. In essence, the problem they describe stems from the fact that a treatment effect heterogeneity has no reason to be independent, or to have null covariance, with a model error term. In our case, fixed effects are hypothesized to be fixed, and deviations from these fixed effects go to the residual term. But the De Chaisemartin and d'Haultfoeuille results indicates that AKM estimates might not be robust to a specification error in this regard. For instance, firm premiums might be sensitive to short term economic fluctuations, as well as workers propensity to change employer. In this case, estimated fixed effects might be systematically biased relative to the firm average premium. While De Chaisemartin and d'Haultfoeuille propose a corrected estimator in their setting, and despite a rapidly growing literature, we are not aware of a generalization of the fix to the AKM setting. de Chaisemartin and D'Haultfoeuille (2022) for instance provide a fix in the case of multiple treatments only when such treatments overlap. In our setting, workers do not have multiple employers at the same time. Engbom et al. (2022) and Lachowska et al. (2020) provide some reassurance. They estimate a derived version of AKM models where firm fixed effects are interacted with years, and find these yearly fixed effects to be highly correlated in time for each firm. Moreover, even if the fluctuations in yearly firm fixed effects are correlated, their results do not depart much from the AKM results, suggesting this kind of heterogeneity, at least, may not have too strong an impact.

3. We also checked with a model inspired by Card et al. (2018), where log-wages are not centered, years are right-hand explanatory variables and age is included as a cubic polynomial constrained to be flat at 40 to avoid colinearity. Results are robust to this change in specification

2.2 Log-wage variance decomposition

Following Card et al. (2013) and Song et al. (2019), we take $V(y) = Var(y_{it})$ as a measure of wage inequalities and observe its evolution through 3 five-years periods : 2002-2006, 2007-2011 and 2012-2006. Ignoring for simplicity of exposition the time-varying workers variables x_{it}, we can describe for each period a decomposition of $V(y)$ as a sum of the estimated variances of θ, ψ, u, and their respective covariances :

$$V(y) = V(\theta) + V(\psi) + V(u) + 2Cov(\theta, \psi) \tag{2}$$

Song et al. further distinguishe within-firms and between-firms components of wage variance, and extend the classic variance decomposition to :

$$V(y) = \underbrace{V(\bar{y}_j)}_{\text{Between-firm component}} + \underbrace{\sum_j m_j \times V(y_i | i \in j)}_{\text{Within-firm component}} \tag{3}$$

$$V(y) = \underbrace{V(\psi) + 2Cov(\bar{\theta}_j, \psi) + V(\bar{\theta}_j)}_{\text{Between-firm component}} + \underbrace{V(\theta_i - \bar{\theta}_j) + V(u)}_{\text{Within-firm component}} \tag{4}$$

With $\bar{y}_j = \bar{y}_{j(i,t)}$ and $\bar{\theta}_j = \bar{\theta}_{j(i,t)}$ the respective expectations on i, t in firm j. By hypothesis the analogous $\bar{u}_j$ is equal to 0. Moments of the distribution of firm variables are weighted by the number of observations per firm m_j. Our interest lies first with the evolution of the sorting component of this decomposition, $Cov(\theta, \phi)$, which is by construction entirely contained in the between-firm component of wage variance.

2.3 Limited mobility bias

We follow Kline et al. (2020) for the description of this bias. They provide a simple framework that neatly generalizes on any quadratic form of the estimated parameters. We start with a simplified notation of our linear model :

$$y_i = z_i' \alpha + u_i \tag{5}$$

With $\alpha = (\beta, \theta, \psi)$ our parameter vector of length $k = 2 + N + J$ and z_i the non-random regressors vector of the (worker * year) i's observation characteristics, including the indicator vector for worker and firm. We note $S_{zz} = \sum_{i=1}^{N^*} z_i z_i'$ the design matrix (with full rank when we limit the sample to the main connected set). Our objects of interest are (weighted) variances and covariances of parts of the α vector and can be described as quadratic forms $\omega = \alpha' A \alpha$ for a chosen symetric matrix $A \in \mathbf{R}^{k \times k}$. We can choose A so as to compute the quantities studied here[4].

Our naive plug-in estimator for ω is thus $\hat{\omega}^{PI} = \hat{\alpha}' A \hat{\alpha}$ with $\hat{\alpha}$ an OLS estimate $\hat{\alpha} = {S_{zz}}^{-1} \sum_{i=1}^{N^*} z_i y_i = \alpha + {S_{zz}}^{-1} \sum_{i=1}^{N^*} z_i u_i$. The estimation error in $\hat{\alpha}$ will result in a systematic bias in $\hat{\omega}^{PI}$ equal to a linear combination of the unknown and possibly heteroscedastic variances σ_i^2 of the error terms u_i. From classic results on quadratic forms, Kline et al. deduce :

$$\mathbf{E}[\hat{\omega}^{PI}] - \omega = trace(A\mathbf{V}[\hat{\alpha}]) = \sum_{i=1}^{N^*} B_{ii} \sigma_i^2 \tag{6}$$

4. For instance, var(WFE) is computed with a matrix A filled with 0 except for the $N \times N$ square corresponding to the N WFE estimates of the parameter vector, which we fill with a generic term $-1/N$ and a diagonal term $1 - 1/N$. Covariances objects are built with analogous matrices, with for instance a $1 - m_j/N$ term for the $(2 + i, 2 + N + j)$ and $(2 + N + j, 2 + i)$ positions in the matrix if the worker i is in firm j and a generic $-m_j/N$ term if the worker i is not working in firm j

With $B_{ii} = z_i' S_{zz}^{-1} A S_{zz}^{-1} z_i$ representing the influence of each (squared) error term on the plug-in estimator. This bias exists for all linear models, but usually for a small parameter dimension k the S^{-2} term insures relatively fast convergence. Here however k is large, and so is, potentially, B_{ii}. Moreover the complex structure of the design matrix, reflecting the complex network of worker / firm connections, is present both in matrices S and A when computing $cov(\theta, \psi)$, leaving way to even stronger bias.

This expression for the bias points to an obvious correction strategy : estimating the σ_i^2 error terms, and thus the bias itself. This can be done with a leave-one-out strategy that is computationally costly, adding a factor of the order $N^* = NT$ to the computation. Kline et al. provide a more tractable estimation method through a high number of random projections (in the hundreds). Bonhomme et al. (2020) still finds the method demanding and further approximate it, though they worry the succession of approximate estimations (with those usual in AKM models) might have consequences that are not well understood.

We favour a split-sampling strategy that only demands two estimations on two half-samples, at worst doubling computing time.

We also implemented Bonhomme et al. (2017) and Bonhomme et al. (2019) strategies in two ways. First, we ran a firm-clustering algorithm with 10 clusters before estimating AKM on firm clusters (rather than individual firms), with the hypothesis that firms fixed effects are discretely distributed with a small number of values. The mobility network between clusters is very dense and each cluster's fixed effect estimate has very low variance, thus correcting the limited mobility bias. The clustering algorithm is a kmeans clustering based on quantiles of the wage distribution, as the identification of clusters can not rely on firm mean wage and must use higher moments of the distribution of wages. Even so, it remains plausible that the segregation of workers could bias the clustering, with firms being clustered based on some combination of their own fixed effects and their average workers' fixed effects. An AKM estimation following this procedure would then show higher sorting, and lower cluster effect variance, than is really the case. Bonhomme et al. (2019) aknowledge the risk and provide in-depth robustness analysis that suggests it is of limited impact in practice.

Second, we followed Bonhomme et al. (2019) in estimating random effects for workers to allow for interactions between workers' types and firm clusters in the model of wages. Similarly to their results, we found limited signs of interaction between firm clusters and workers types : higher wages firms tend to give higher premiums to higher types workers than to other workers. The overall weight of these interactions appears limited however.

2.4 Bias correction with split-sampling estimation of quadratic forms

For exposition, we first suppose that observations are randomly split in two half samples, and that each sample retains the same connectivity as the full sample. Our split-sampling plug-in estimator for the quadratic form ω becomes $\hat{\omega}^{SP} = \hat{\alpha}_0' A \hat{\alpha}_1$ with $\hat{\alpha}_s$ an OLS estimate in the sample $I_s, s = 0, 1$ of size N_s : $\hat{\alpha}_s = S_{zz,s}^{-1} \sum_{i \in s}^{N_s} z_i y_i = \alpha + S_{zz,s}^{-1} \sum_{i \in s}^{N_s} z_i u_i$ that we can express as $\hat{\alpha}_s = \alpha + \epsilon_s$. We calculate the bias by expressing a scalar as the trace of a (1,1) matrix, as in the classic demonstration on the expectation of quadratic forms

$$\begin{aligned}
\mathbf{E}[\hat{\omega}^{SP}] &= \mathbf{E}[\hat{\alpha}_0' A \hat{\alpha}_1] = \mathbf{E}[trace(\hat{\alpha}_0' A \hat{\alpha}_1)] & \\
&= \mathbf{E}[trace((A\hat{\alpha}_0)' \hat{\alpha}_1)] & \text{by symmetry of A} \\
&= \mathbf{E}[trace(A\hat{\alpha}_0 \hat{\alpha}_1)'] & \text{by propriety of trace()} \\
&= trace(A\mathbf{E}[\hat{\alpha}_1 \hat{\alpha}_0']) & \text{A non-random, trace() is linear}
\end{aligned}$$

We further have :

$$\begin{aligned}\mathbf{E}[\hat{\alpha}_1\hat{\alpha_0}'] &= \mathbf{E}[(\alpha+\epsilon_1)(\alpha+\epsilon_0)'] \\ &= \alpha\alpha' + \mathbf{E}[\epsilon_1\epsilon_0']\end{aligned}$$

And :

$$trace(A\alpha\alpha') = \omega$$

The bias is thus equal to :

$$\mathbf{E}[\hat{\omega}^{SP}] - \omega = trace(A\mathbf{E}[(S_{zz,1}^{-1}\sum_{i\in I_1}^{N_1} z_i u_i)(S_{zz,0}^{-1}\sum_{j\in I_0}^{N_0} z_j u_j)'])$$

$$\begin{aligned}\mathbf{E}[\hat{\omega}^{SP}] - \omega &= trace(A\mathbf{E}[(S_{zz,1}^{-1}\sum_{i\in I_1}^{N_1} u_i z_i)(S_{zz,0}^{-1}\sum_{j\in I_0}^{N_0} u_j z_j)']) \\ &= trace(AS_{zz,1}^{-1}\mathbf{E}[\underbrace{(\sum_{i\in I_1}^{N_1} u_i z_i)(\sum_{j\in I_0}^{N_0} u_j z_j)'}_{\text{matrix } (b_{lm})}](S_{zz,0}^{-1})')\end{aligned}$$

with generic term :

$$b_{lm} = \sum_{i\in I_1}^{N_1} u_i z_{l,i} \sum_{j\in I_0}^{N_0} u_j z_{m,j}$$

This term has null expectation under mild conditions : 1. null conditional expectation $\mathbf{E}[u|z] = 0$ and 2. independence of $u_i, i \in I_1$ and $u_j, j \in I_0$. If the variance-covariance matrix of u is diagonal, the bias disappears whatever the matrices A and S_i might be. The second condition might be violated, if for instance u_i are correlated for different years of the same employer / employee pair, which is likely. Of course there is a cost in increased uncertainty coming from the reduced effective sample size ($S_{zz,s}$ has half the observations of S_{zz}). With our data, we observe that this uncertainty is small compared to the size of the bias reduction effect. One can reduce this uncertainty by repeatedly estimating the quadratic form through split-sampling and averaging the results. The procedure reaches arbitrary precision, only limited by the computational cost. On the small sample, we show the average of 20 split-sampling procedures (40 estimates) and the corresponding 95% confidence interval showing the size of the remaining sampling uncertainty.

The main limitation of split sampling is the impact of the split on the bipartite graph and its main connected set, and the sample-splitting strategy has to be considered in this regard. In each split sample, the main connected set is smaller than in the original sample and both are distinct, so that the common sample of workers and firms belonging to the main connected set in both split samples is reduced, and so is the corresponding parameter vector of individual effects.

The most simple split strategy is a direct random split of observations in two equally sized samples. By balancing the sampling by worker, splitting for each worker the periods of observation, one increases the odds that each worker is present in both samples' main connected set. We dub this method "period splitting". On the contrary, by splitting individuals rather than observations, one increases the connectivity in each set (because individual careers are kept intact), but each worker's fixed effect is estimated only once : one looses the capacity to correct the $var(WFE)$ quadratic form through split sampling. If this splitting of individuals is balanced by firm, it increases the odds that each firm's fixed effect is estimated in each sample. We dub this method "firm splitting" [5].

5. Chanut (2018) provides an analysis of this split sampling strategy to correct the limited mobility

Period splitting might not completely correct the limited mobility bias for the reason mentioned above : it is likely that u_i are correlated for several observations of the same employer / employee pair. This problem is attenuated by the specifics of our setting. Because we keep only observations with full year jobs, movers are generally observed only four years or less among the five in the panel. It is unlikely that after the random split of these observations, an individual worker would remain a mover in both samples. Because the estimation relies exclusively on movers, a given individual residuals would generally not be correlated to errors on both sides of the split. Still, by keeping all observations of one individual on one side of the split, the firm splitting method avoids entirely this drawback. Consequently, we favor firm splitting to compute a debiased sorting effect, but use period splitting to compute the complete variance decomposition.

2.5 Explaining the rise in sorting through covariance decomposition

Decomposition of the covariance

Covariance decomposition is similar to variance decomposition. Say we want to understand the links between this covariance and k other variables $X = (X_1, ..., X_j, ..., X_k)$, such as sex, occupation, etc. Start with a general formula derived from the definition of covariance :

$$Cov(WFE, FFE) = \mathbf{E}[Cov(WFE, FFE|X)] + Cov[\mathbf{E}(WFE|X), \mathbf{E}(FFE|X)] \tag{7}$$

The first term is the "unexplained" (by X) covariance, the second term the "explained" covariance. When $k > 1$, the interpretation of the explained term decomposition draws on the interactions between the variables. Empirical estimation is straightforward for categorical explanatory variables, but more challenging for continuous variables : one has to approximate the $\mathbf{E}(FFE|X)$, $\mathbf{E}(WFE|X)$ and $Cov(WFE, FFE|X)$ functions of X. We implement the decomposition at the firm level, where any categorical variable at the worker level (such as sex) becomes a quantitative variable or variables at the firm level (such as the share of women in the workforce).

Linear models are natural choice. Let's assume that :

$$WFE = \alpha_w + X\beta_w + \epsilon_w$$
$$FFE = \alpha_f + X\beta_f + \epsilon_f$$

With $Cov(\epsilon_w, X) = Cov(\epsilon_f, X) = 0$. The covariance can then be decomposed as :

$$\begin{aligned} Cov(WFE, FFE) &= Cov(X\beta_w + \epsilon_w, X\beta_f + \epsilon_f) \\ &= \underbrace{Cov(X\beta_w, X\beta_f)}_{\text{Explained cov}} + \underbrace{Cov(\epsilon_w, \epsilon_f)}_{\text{Unexplained cov}} + \underbrace{Cov(X\beta_w, \epsilon_f) + Cov(\epsilon_w, X\beta_f)}_{\text{Equals 0}} \end{aligned}$$

The sample covariance can be computed as an empirical mean. WFE and FFE are centered in the sample. We further assume $X'X$ is invertible and define $\tilde{X} = (X'X)^{-1}X'$ so that $\hat{\beta}_w = \tilde{X}W$

bias. He describes a way to compute such a split, uses this method on the French narrow panel and shows, on a toy example, that it succeeds in correcting the bias.

and $\hat{\beta}_f = \tilde{X}F$.

$$\widehat{Cov}(WFE, FFE) = 1/n \sum_i WFE_i FFE_i \tag{8}$$

$$= 1/n \sum_i \underbrace{\hat{\beta}_f(X_i - \overline{X})(X_i - \overline{X})'\hat{\beta}_w}_{\text{Explained}} + \underbrace{\hat{\epsilon}_{f,i}\hat{\epsilon}_{w,i}}_{\text{Unexplained}} + \underbrace{(X_i - \overline{X})'(\hat{\beta}_f\hat{\epsilon}_{w,i} + \hat{\beta}_w\hat{\epsilon}_{f,i})}_{\text{0 in expectation}} \tag{9}$$

Using the decomposition to understand the evolution of sorting

We have two additional requirements for the decomposition of covariance to help us understand the dynamics of sorting. First, it must be used to compare covariance between two periods. Second, the explanatory variables need have a plausible causal interpretation. We develop a new idea. Consider firms j : we pool estimates for period 1 and 3 together, weight by the number of workers' observations per firm and period, and build firm level explanatory variables. We can regress worker fixed effects and firm fixed effects on :

$$WFE_{jt} = \alpha_w + W_j + \beta_w * X_{jt} + \epsilon_{jt}$$
$$FFE_{jt} = \alpha_f + F_j + \beta_f * X_{jt} + \omega_{jt}$$

where W_j and F_j capture "meta" firm fixed effects in both regressions, respectively, to account for selection in unobservables. In these regressions, the β parameters capture the link for a given firm between a change in, say the size of a firm between period 1 and 3, and the change in the firm's fixed effect and the firm's average worker fixed effect, respectively. We then consider each period independently.

We can decompose covariance for period 1 :

$$\begin{aligned} Cov(WFE_1, FFE_1) = Cov(\alpha_w + W_1 + \beta_w * X_1 + \epsilon_1, \alpha_f + F_1 + \beta_f * X_1 + \omega_1) = \\ \boldsymbol{Cov(W_1, F_1)} + Cov(W_1, \beta_f * X_1) + Cov(W_1, \omega_1) + \\ Cov(\beta_w * X_1, F_1) + \boldsymbol{Cov(\beta_w * X_1, \beta_f * X_1)} + Cov(\beta_w * X_1, \omega_1) + \\ Cov(\epsilon_1, F_1) + Cov(\epsilon_1, \beta_f * X_1) + \boldsymbol{Cov(\epsilon_1, \omega_1)} \end{aligned}$$

We can write the same for period 3. It follows that the difference in the covariance between period 3 and 1 can be expressed as the difference between the 9 components. How does one understand this decomposition? Beginning with the terms in bold, the "diagonal" terms when the decomposition is shown in matrix form, we have the direct effects of selection (of firms), covariates and residuals. Here, because we estimate on both periods and then decompose on each period separately, the residual do not disappear. The first diagonal term represents pure selection : the population of firms in the first period might have a stronger, or weaker, covariance between W_j and F_j than firms in the last period. This selection is driven by singleton firms (firms observed only in the first or last period) but also by a change in firm size (and hence weights). The diagonal residual term reflects systematic changes in unobservables between the to periods. Finally the second diagonal term, between covariates, can itself be decomposed by covariate into diagonal and interaction terms. The diagonal terms reflect the variance of the explanatory variable (in each period), multiplied by the product of the parameters β in each regression. The interaction terms similarly are the product of the covariance of the two variables (in each period) multiplied by the product of their respective parameters.

3 Results

3.1 A robust rise in sorting

In table 1, these two decompositions 2 and 4 are applied to our three periods, where WFE and FFE stand for "workers fixed effects" and "firms fixed effects". We use the classic AKM estimates of θ and ψ fixed effects as our baseline estimates, before applying any of the corrections discussed above. Overall wage inequalities in the private sector, measured as log-hourly wage variance, went down slightly between 2002 and 2016 in France. The decomposition of the variance shows that while the within-firm log-hourly wage variance was also diminishing, between-firm wage inequalities rose during the same period. They accounted for 42% of total log-hourly wage variance in 2002-2006, and 48% in 2012-2016. Log-hourly wage variance was lower by 3.7% in the third period compared to the first : within-firm inequalities accounted for 204% of this evolution, and between-firm inequalities for -104%. France's diminishing wages inequalities during the period are atypical among developed countries, but the rise of between-firm wage inequalities matches the results of Song et al.. for the US from 1978 to 2013. It suggests that the same mechanisms that associate firm dynamics and growing wage inequalities in most rich countries operate in France.

This rise is robust to the various checks we conducted, most notably to the correction of the limited mobility bias, which appear very important, even when restricting to firms with more than 20 observations per year. All results are presented in figure 1. The lowest curve shows uncorrected sorting as measured on the small panel, the very data used in the original Abowd et al. (1999) paper, but for subsequent years.

The intensity of the rise is lower in corrected estimates, suggesting the limited mobility bias decreased overall on the period, possibly due to increased connectivity of the firm-workers network. This evolution shows that a perfect stability of the limited mobility bias cannot be assumed, and that the dynamics of sorting must be measured based on corrected estimates.

Split sampling correction behaves as expected, with period-splitting showing signs of an incomplete correction of the bias compared to firm-splitting. The firm-clustering method gives results reasonably close to the split sampling, reproducing results from Bonhomme et al. (2020) when comparing their clustering and random effect model to the Kline et al. (2020) leave one out method. This might be indicative of a small upward bias in the firm clustering method, as previously discussed. Alternatively, it might come from uncomplete correction of the bias from split-sampling, even with firm-splitting, if for instance correlated moves and residuals of different workers from the same firm (which is the case in merges and acquisitions) weight on the estimation.

Long-term series are imperfect, as is clear when compared to exhaustive data estimates on recent years, but they might provide some indication of past trends. They suggest that most of the rise in sorting actually predates our main period of study, and show fluctuations that are evocative of pro-cyclicality.

3.2 A central role for occupations

The covariance decomposition described in section 2.5 leads to some insights into the dynamics of the rise in sorting. First, the firms (meta-) fixed effects covariance explains an important share of the rise (table **??**). Since these meta-fixed effects are fixed, identical in period 1 and 3, this conveys a selection effect : firms that contribute to sorting, for reasons that are not captured by observed covariates, are more numerous and bigger in 2012-2016. More precisely, we observe that employment grew mostly in two types of firms : firms with high premium and high wage workers present both in period 1 and period 3, and firms with low premium and low-wage workers observed only in period 3 (new firms) as opposed to similar firms observed only in period 1.

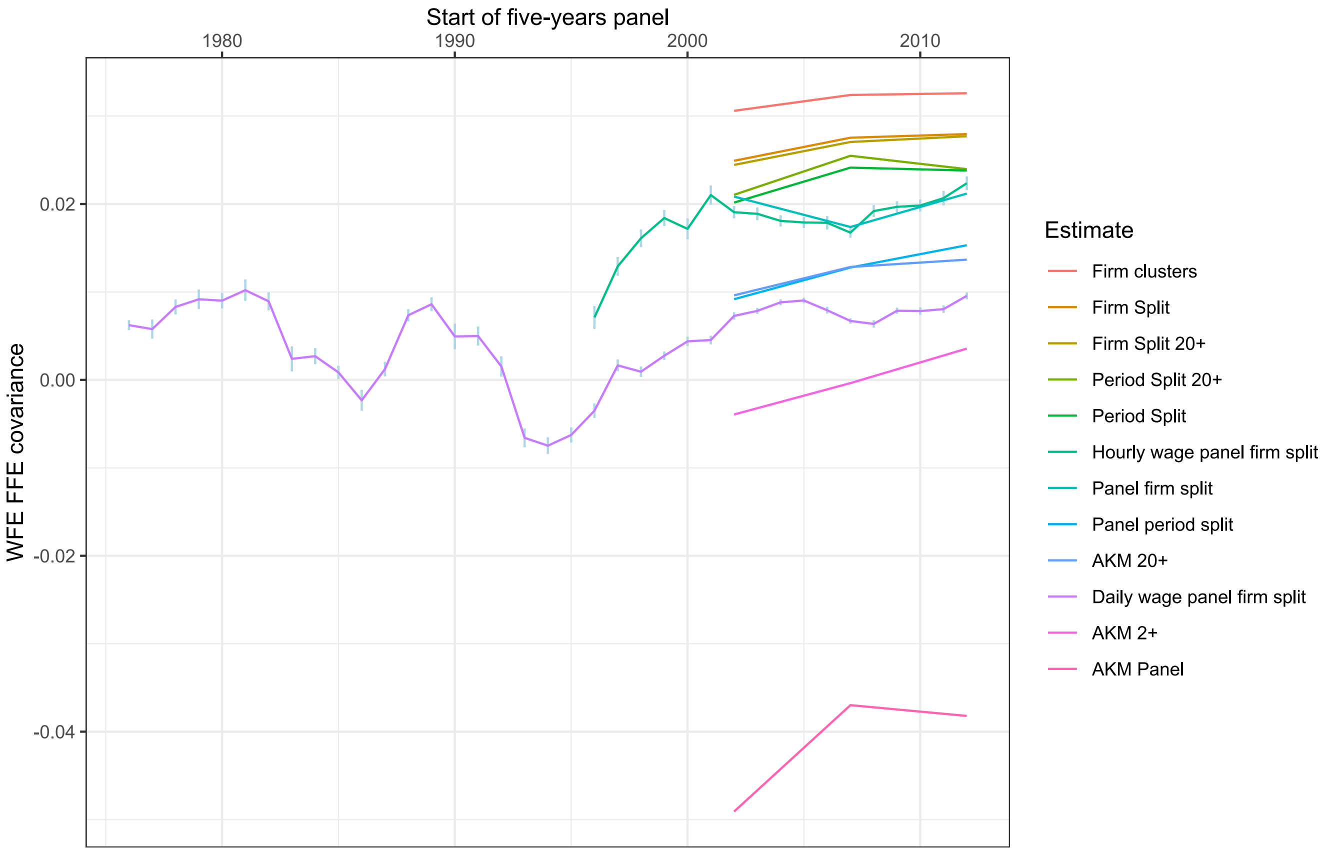
Table 1 – Decomposition of wage variance and its evolution

		2002-2006		2007-2011		2012-2016		Change from 2002-2006 to 2012-2016	
		Comp.	Share	Comp.	Share	Comp.	Share	Comp.	Share
Total variance	Var(y)	0,214		0,211		0,206		-0,008	
Components of variance	Var (WFE)	0,165	77,1	0,166	78,8	0,158	76,6	-0,007	90,2
	Var (FFE)	0,030	14,0	0,029	14,0	0,025	12,0	-0,005	65,2
	Var(Xb)	0,024	11,1	0,034	16,0	0,016	7,8	-0,008	97,4
	Var(u)	0,009	4,1	0,008	4,0	0,008	4,0	-0,001	6,8
	2*Cov(WFE,FFE)	-0,004	-1,8	0,000	-0,2	0,005	2,2	0,008	-106,4
	2*Cov(WFE,Xb)	-0,012	-5,7	-0,029	-13,7	-0,007	-3,2	0,006	-69,5
	2*Cov(FFE,Xb)	0,002	1,2	0,002	1,1	0,001	0,6	-0,001	16,3
Between-firm variance	Var(m_y)	0,091	42,2	0,095	45,2	0,099	47,9	0,008	-103,9
	Var (m_WFE)	0,058	27,2	0,063	29,7	0,066	32,0	0,008	-97,3
	Var (FFE)	0,030	14,0	0,029	14,0	0,025	12,0	-0,005	65,2
	Var(m_Xb)	0,004	1,8	0,006	2,6	0,003	1,3	-0,001	16,6
	2*Cov(m_WFE,FFE)	-0,004	-1,8	0,000	-0,2	0,005	2,2	0,008	-106,4
	2*Cov(m_WFE,m_Xb)	0,000	-0,1	-0,004	-2,0	0,000	-0,2	0,000	1,6
	2*Cov(FFE,m_Xb)	0,002	1,2	0,002	1,1	0,001	0,6	-0,001	16,3
Within-firm variance	Var(diff_y)	0,124	57,8	0,115	54,8	0,108	52,1	-0,016	203,9
	Var (diff_WFE)	0,107	49,9	0,103	49,1	0,092	44,6	-0,015	187,4
	Var(diff_Xb)	0,020	9,3	0,028	13,4	0,014	6,5	-0,006	80,8
	Var(u)	0,009	4,1	0,008	4,0	0,007	3,3	-0,002	26,8
	2*Cov(diff_WFE,diff_Xb)	-0,012	-5,6	-0,025	-11,7	-0,006	-3,0	0,006	-71,1
	2*Cov(diff_WFE,u)	0,000	0,0	0,000	0,0	0,000	0,0	0,000	0,0
	2*Cov(diff_Xb,u)	0,000	0,0	0,000	0,0	0,000	0,0	0,000	0,0
Segregation Index	$\frac{Var(m\_WFE)}{Var(WFE)}$	0,351		0,379		0,419		0,068	
N* (main connected set)		41 703 340		44 733 304		47 038 310			

Var(y) : variance of the natural log of hourly wage, Var(WFE) : variance of worker fixed effects, Var(FFE) : variance of firm fixed effects, Var(Xb) : variance of covariates, Var(u) : variance of residual.
All firms and individuals in firms with at least 2 employees are included. Only individuals employed for at least 360 days by the same firm during the year are included for a given year. Individuals and firms in public administration are not included.

Secondly, when focusing on diagonal terms, capturing the most direct effect of each variable, we find that among explanatory variables we observe, the occupational structure appears to play a central role. Occupations linked to both high (respectively low) workers and firm effects have become more strongly associated with time (table **??**). On the other hand, neither age or gender segregation, or size an added value variance between firms appear to play a noticeable role in the rise in sorting.

FIGURE 1 – Sorting, all estimates



Sorting estimates by five years periods. "Panel" designates the small panel. Long term series are computed on the small panel, on rolling 5 years period, corrected by firm-splitting. Mean estimate and confidence intervals computed on repeated (split) sampling with 20 repetitions

TABLE 2 – Decomposition of the difference in sorting between period 3 and 1

	FFE regression						
WFE regression	Residuals	FE	VA	Size	% Female	Age	Occupation
Residuals	**7.90%**	49.24%	5.12%	-0.21%	-0.57%	1.32%	6.51%
Fixed Effects	-30.76%	**30.91%**	7.31%	6.80%	-1.45%	-0.83%	32.99%
Value Added	-4.48%	5.05%	**1.00%**	-0.01%	0.73%	-0.16%	4.41%
Size	0.13%	0.02%	0.00%	**0.56%**	0.14%	-0.04%	2.70%
% Female	1.64%	3.20%	1.18%	0.35%	**0.26%**	-0.34%	4.19%
Avg Age	2.11%	13.15%	0.87%	0.35%	1.12%	**-0.25%**	1.33%
Occupation	-54.72%	-22.40%	7.48%	-5.60%	4.14%	-0.48%	**18.08%**

All firms and individuals in firms with at least 20 employees are included. Only individuals employed for at least 360 days by the same firm during the year are included for a given year. Individuals and firms in public administration are not included. VA stands for value added per worker.
Values computed as explained in section 2.5. All covariates are an average for that firms across the years of the period considered.

Such a decomposition can be hard to interpret, because of interactions between observed variables, and between observed variables and unobservables (residuals). Attempts at synthesizing remains dependant on the included variables in the baseline model, and the order in which variables are added, in a sequential approach. More sophisticated methods to summarize the impact of a given variable on variance could plausibly be adapted to our setting, while being computationally costly. In table 3, we show the results of two simpler methods. First, starting with a base model with only the meta-fixed effects, we look at the impact that each variable has, when added to this base model, on the share of the residuals covariance in the rise of sorting. Alternatively, starting from the complete model as a base model, we look at the impact that each variable has, when ommitted, on the same covariance of residuals. Both methods point to the reduction in the unexplained part of the rise that each variable provide : occupations are the only variable with an impact.

TABLE 3 – **Variance of Residuals (%)**

	"Addition"	*"Omission"*
Meta-FE (Benchmark)	8.29%	
Value-Added	8.21%	7.81%
Occupation	7.67%	8.26%
Incidence of Female	8.30%	7.89%
Avg. Age	8.44%	7.82%
Size	8.25%	7.91%

4 Discussion

The AKM model has proven a robust description of wages. But the limited mobility bias is a serious limitation that led to an important underestimation of sorting. Our results confirm that, once corrected for this bias, sorting accounts for more than 10% of overall wage inequalities, measured as the variance of the log-wage. Although less seriously, the bias also impacts the measure of the evolution of sorting, likely because mobility intensity and patterns do evolve in time. We found however that sorting did increase in France, as it did in the USA and Germany, even though log-wage variance in France remained stable on the period. Split-sampling is an easy patch, provably correct under reasonable hypothesis, and suggests that clustering methods might suffer from a slight upward bias.

The causes of this rise in sorting are more elusive. Like measures of inequalities, sorting is a distributional statistic, not a characteristic of individuals or firms. It is not directly amenable to classical econometric analysis. Expanding variance decomposition methods to covariance decomposition, we found suggestive evidence that sorting in workers is associated to sorting in occupations : an increase in occupational segregation in firms along the occupational hierarchy. We also find that firm demographics explain a large part of the rise in sorting : high-premiums, high wages firms have grown more that others, and newly created firms tend to be more often low-wage, low-premium than the one they replace. Both phenomenon would point toward a structural evolution in the division of work between firms, such as an increased externalisation of low-value added tasks. Other statistical sources might better inform these phenomenon.

Other important open questions are both methodological and substantive. Methodologically, there are additional limitations that are not yet well understood. One is that fixed effects are not fixed. The complete consequences of this specification error are difficult to grasp for the moment, but they might impact the measure of sorting, as well as the other component of the decomposition, especially when short-term economic fluctuations are large. A linked question are the measures of the age and experience components of wages, when it matters to disentangle yearly effects, age and individual effects. We found fluctuations in the interactions between these terms that are suggestive of some estimation artefact, but still resistant to alternative specifications.

On the substance, there is more to explore about the interplay of French institutional features and sorting. It appears that wage inequalities in France have been controlled, for most of the period, by an increase in the redistributive power of payroll taxes and by irregular increments in the minimum wage. Both phenomenon necessarily impact the shape of the distribution of wages and interact with firm pay policies, the job market, and sorting. There is yet more to be learned from the French case.

Bibliographie

Références

Abowd, J. M., F. Kramarz, P. Lengermann, and S. Pérez-Duarte (2004) : "Are good workers employed by good firms ? A test of a simple assortative matching model for France and the United States," *Unpublished Manuscript.*

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999) : "High wage workers and high wage firms," *Econometrica*, 67, 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008) : "High wage workers and low wage firms : negative assortative matching or limited mobility bias ?" *Journal of the Royal Statistical Society : Series A (Statistics in Society)*, 171, 673–697.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020) : "How Much Should we Trust Estimates of Firm Effects and Worker Sorting ?" Tech. rep., National Bureau of Economic Research.

Bonhomme, S., T. Lamadon, and E. Manresa (2017) : "Discretizing unobserved heterogeneity," *University of Chicago, Becker Friedman Institute for Economics Working Paper.*

——— (2019) : "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

Borovičková, K. and R. Shimer (2017) : "High wage workers work for high wage firms," Tech. rep., National Bureau of Economic Research.

Bozio, A., T. Breda, and M. Guillot (2020) : "The Contribution of Payroll Taxation to Wage Inequality in France," .

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018) : "Firms and labor market inequality : Evidence and some theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., J. Heining, and P. Kline (2013) : "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 128, 967–1015.

Chanut, N. (2018) : "Distinguishing Between Signal and Noise in the Measurement of the Firm Wage Premium," *Available at SSRN 3470571.*

De Chaisemartin, C. and X. d'Haultfoeuille (2020) : "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–96.

de Chaisemartin, C. and X. D'Haultfoeuille (2022) : "Two-way Fixed Effects and Differences-in-Differences Estimators with Several Treatments," Tech. rep., National Bureau of Economic Research.

Drenik, A., S. Jäger, M. P. Plotkin, and B. Schoefer (2020) : "Paying outsourced labor : Direct evidence from linked temp agency-worker-client data," Tech. rep., National Bureau of Economic Research.

Engbom, N., C. Moser, and J. Sauermann (2022) : "Firm pay dynamics," Tech. rep., National Bureau of Economic Research.

Gerard, F., L. Lagos, E. Severnini, and D. Card (2018) : "Assortative matching or exclusionary hiring ? The impact of firm policies on racial wage differences in Brazil," Tech. rep., National Bureau of Economic Research.

Godechot, O., P. Apascaritei, I. Boza, L. F. Henriksen, A. S. Hermansen, F. Hou, N. Kodama, A. Křížková, J. Jung, M. M. Elvira, et al. (2020) : "The great separation : Top earner segregation at work in high-income countries," Tech. rep., MaxPo Discussion Paper.

Goldschmidt, D. and J. F. Schmieder (2017) : "The rise of domestic outsourcing and the evolution of the German wage structure," *The Quarterly Journal of Economics*, 132, 1165–1217.

Jochmans, K. and M. Weidner (2019) : "Fixed-Effect Regressions on Network Data," *Econometrica*, 87, 1543–1560.

Kline, P., R. Saggio, and M. Sølvsten (2020) : "Leave-out estimation of variance components," *Econometrica*, 88, 1859–1898.

Lachowska, M., A. Mas, R. D. Saggio, and S. A. Woodbury (2020) : "Do firm effects drift ? Evidence from Washington administrative data," Tech. rep., National Bureau of Economic Research.

Schoefer, B. and O. Ziv (2021) : "Productivity, Place, and Plants : Revisiting the Measurement," Tech. rep., CEPR Discussion Paper No. DP15676.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019) : "Firming up inequality," *The Quarterly journal of economics*, 134, 1–50.

Tomaskovic-Devey, D., A. Rainey, D. Avent-Holt, N. Bandelj, I. Boza, D. Cort, O. Godechot, G. Hajdu, M. Hällsten, L. F. Henriksen, et al. (2020) : "Rising between-workplace inequalities in high-income countries," *Proceedings of the National Academy of Sciences*, 117, 9277–9283.