Stuck in the middle? Occupation-specific commute-wage trade-off at the metropolitan level

Maxime Liégey *   Nathalie Picard †

February 15, 2024

Preliminary and incomplete, please do not cite

Abstract

How do middle-skilled workers trade off wages against commuting time, as compared to high- and low-skilled ones? In this paper, we leverage a quasi-exhaustive panel of jobs in France, to explore how unobserved heterogeneity can help characterize workers' trade-off at the metropolitan area level. We use estimated worker- and employer fixed effects in order to construct two measures of how constrained workers are, depending on their broad occupational group. A first measure, the Commute-Wage Gradient (CWG), captures the trade-off that a marginal entrant into a metropolitan area would face, while a second, labelled Monopsony Power Measure (MPM), gauges the extent to which employers' monopsony power influences workers' outside options at alternative employers. We find that middle-skilled workers are *stuck*, in the sense that higher-wage *middle*-skilled earners have to commute more, than otherwise identical high- or low-skilled workers, while being at the same time more subject to monospony power. By contrast, high-skilled workers are less constrained according to both measures. Last, we document non-monotonicites in the case of low-skilled workers, which is hardly consistent with a job ladder model.

*Université de Strasbourg: liegeym@gmail.com

†Université de Cercy and Université de Strasbourg: picardn@unistra.fr

The financial support of the ANR TOPAZE (ANR-18-CE26-0009) is gratefully acknowledged. The authors thank participants at TOPAZE Workshops organized at CYU, on March 9^{th} and June 22-23^{th} of 2023, as well as participants at the ITEA Annual Conference in Santander, Spain, in June 2023. The technology to access the data remotely is supported by a public grant overseen by the French National Research Agency (ANR) as part of the Investissements d'avenir program (ANR-10-EQPX-17 – Centre d'accès sécurisé aux données, CASD).

1 Introduction

How do high-, middle- and low-skilled workers trade off wage against commuting time, depending on the distribution of job opportunities in a given metropolitan area? On the one hand, geographical, historical and institutional factors will to some degree determine the distribution of jobs and impact spatial residential sorting,[1] so that city structure can to some extent be considered as exogenous, at least over the short to medium-run. On the other, the literature has already identified key determinants of the wage-commuting trade-off, namely: gender,[2] occupation,[3] the life cycle,[4] and employer's industry.[5] Yet, the role of unobserved heterogeneity in describing workers' commute-wage trade-off has been largely ignored in the literature. We aim here, leveraging a new, quasi-exhaustive matched employer-employee data in France, at exploring how unobserved heterogeneity can help characterize workers' commute-wage trade-off at the metropolitan level, through the lens of a canonical job ladder model.[6]

The literature on the determinants of wage dispersion has indeed long acknowledged the importance of unobserved heterogeneity, as captured by time-invariant worker- and employer fixed effects, in accounting for the wage variance (Abowd et al. (1999), Bonhomme et al. (2023)). We study here the role of unobserved heterogeneity in the commute-wage trade-off, to determine to what extent higher-wage workers of a given broad occupational group, in a given metropolitan area, tend to be longer-commute workers, as captured by the wage and commuting-time unobserved, time-invariant component.

High-, middle and low-skilled workers are known to face different commute-wage

[1] See Heblich et al. (2021)

[2] See Inoa et al. (2015), Le Barbanchon et al. (2020) and Farré et al. (2022).

[3] See Sang et al. (2011).

[4] See French et al. (2020) and Inoa et al. (2015).

[5] See Gimenez-Nadal et al. (2018).

[6] We use the quasi-exhaustive panel of French workers by implementing the SAS code used by Babet et al. (2023), which enables the concatenation of the exhaustive *Fichiers Postes* from the French DADS (see Olivier Godechot's webpage). The traditional $1/12^{th}$ sample of the *Panel DADS* used so far would pose serious limitations in implementing two-way fixed effect / "AKM" estimations (named after Abowd et al. (1999)) at the metropolitan level.

trade-offs. High-skilled workers (*managers* and *engineers*) typically earn a higher wage income, which allows them to consider a wider set of residential opportunities in a given metropolitan area, while their jobs tend to be clustered geographically, thus inducing them to commute more. By contrast, job opportunities for the middle- and low-skilled workers are more evenly distributed in space, yet these workers typically enjoy a lower bargaining power *vis-à-vis* employers,[7] so their commute-wage trade-off is to a larger extent determined by employers' monopsonistic power.[8]

In this paper, we characterize workers' commute-wage trade-off, depending on their broad occupational category, by proposing two measures, in order to assess how constrained workers are, regarding their wage-commuting trade-off. The first measure, which we label the *Commute-Wage Gradient* (CWG), consists in the elasticity of the estimated commuting-time unobserved component, to the wage unobserved component, as an empirical counterpart of the slope of a representative worker's *indifference curve*, in a log-Commute/Wage plane: the steeper the slope, the longer the commute for higher-wage workers.[9] The second measure, which we label the *Monopsonistic Power Measure* (MPM) aims at capturing employers' market power in shaping workers' commute-wage trade-off. It consists in comparing the observed CWG, with an analogous, counterfactual one, labelled CWG*, where workers of a given broad occupational group are randomly assigned to accessible employers, within a given metropolitan area. The rationale for such a measure comes for the theory of job ladder models,[10] whereby the inverse of the distance between the observed job distribution, and a latent job offer distribution from which workers sample job opportunities at a given job arrival rate, captures the overall strength of employers' market power in shaping the observed commute-wage schedule: the lower

[7] See Cahuc et al. (2006)

[8] We indeed document in the appendix A.3 that labor demand factors, namely the employer fixed effects and the sorting of workers to employers, determine a larger fraction of the wage and commuting time distributions for the middle- and low-skilled workers, than for the high-skilled ones.

[9] Such an indifference curve is upward-sloping, as the commuting time enters negatively in a worker's utility function. This measure captures the overall trade-off that a marginal entrant worker of a given broad occupation, in a given metropolitan area, would face.

[10] Manning (2003)

the job-arrival rate, the larger the employers' market power, and the smaller the distance between the observed job distribution, and the latent job offer one.[11]

In this paper, we first show that middle-skilled workers overall face a more constrained trade-off as measured by a steeper CWG, while being at the same time more subject to monopsonistic power, than both high- and low-skilled workers. In other words, workers in the *middle* of the wage distribution (the middle-skilled) are to some extent "stuck", in the sense, on the one hand, that a marginal middle-skilled entrant into the metropolitan market would face much longer commuting times were she to earn higher wages, and, on the other, in the sense that alternative, utility-improving job offers are relatively scarce. By contrast, we show that high-skilled workers enjoy a less-contrained commute-wage trade-off, as measured by a flatter CWG, while benefiting to a larger extent from utility-improving job mobility, that is, being less subject to monopsonistic power. Finally, we document a U-shaped pattern in low-skilled workers' commute-wage schedule, which can hardly be explained by a job ladder model: lower wage earners among the low-skilled tend to commute more than median wage earners, who in turn commute less than higher wage earners. We interpret this finding by the observation that low-skilled workers tend to reside in the surrounding *exurbs*, where wage income is lower than in principal cities, so that commuters either find lower-paying jobs in the surrounding area, or higher-paying ones at the principal city of the metropolitan area.

The paper is organized as follows. Section 2 presents the data and documents average commuting time differences between broad occupational groups, as well as the role of labor demand factors in determining workers' wage and commuting time. Section 3 describes the two-way fixed effect/AKM[12] estimations at the metropolitan areas, and

[11] To be more specific, Manning (2003) uses a canonical job ladder model to propose the *fraction of hires from non-employment* as a natural proxy for *monosponistic power*. Yet, considering a well-known limitation of French matched employer-employee data (DADS), namely, the absence of information on non-employment spells, we closely follow the spirit of the canonical job ladder model in proposing the distance between the observed job distribution, and the latent job offer distribution, as a proxy for monopsonistic power.

[12] We henceforth use the expressions "two way fixed effect" and "AKM" interchangeably.

defines the two measures characterizing workers' commute-wage trade-off, and presents the results.

2 Wage and commuting patterns by occupational group

2.1 Data

We use the *exhaustive DADS matched employer-employee panel*[13] in order to track workers at the metropolitan area level[14] over a 10-year period, from 2010 to 2019.

We use the same aggregate classification of workers into broad occupational groups, as in Davis et al. (2020),[15] who construct these groups using matches' 2-digit occupational code from the French "*Catégorie Sociale*" (CS) classification.

Individual commuting time between French municipalities (*communes*) is computed using the *Distancier Metric*.[16] In the DADS, each entry records a worker's municipality of residence, as well as the employer's[17] municipality of location, so individual commuting time is the imputed travel time between these two municipalities, by car, off pear hours.

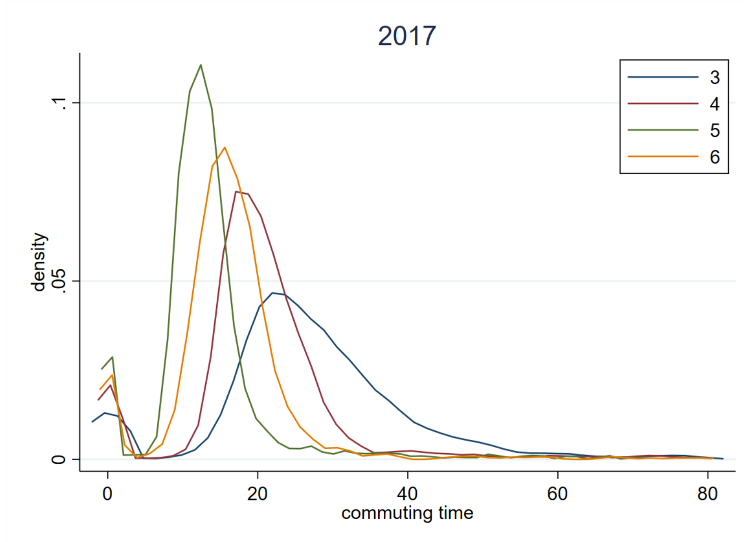


Figure 1: Kernel density estimation of commuting time, by *Catégorie Sociale* (CS), from *Déclaration Annuelle de Données Sociales* (DADS), in 2017

2.2 High-skilled workers commute more than middle-skilled ones, who in turn commute more than low-skilled workers

Figure 1 shows that managers and engineers (*Catégorie Sociale*=3) overall commute significantly more, overall, than intermediate professions (*Catégorie Sociale*=4),[18] who in turn commute more than clerks and workers (*Catégorie Sociale*=5 or 6). This ordering is preserved, when looking at the the share of workers living and working in the same municipality (commuting time = 0): managers and engineers are least likely to live and work in the same municipality, workers and clerks being more likely to do so.

[13]The standard DADS panel used thus far consists in a $1/12^{th}$ sample of French wage earners, which is a serious limitation when it comes to run AKM regressions at the local geographical level. We are grateful to Babet et al. (2023) for sharing their code to construct the exhaustive DADS panel, from the *Fichier Postes*.

[14]By *metropolitan area* we mean a French *Aire d'Attraction Urbaine*, a harmonized classification based upon the *Functional Urban Area* concept defined by Eurostat and OECD.

[15]See table in Figure A.1 in appendix A.1. Following standard practice in the job polarization literature, 2-digit occupations are grouped into "High-/Middle-/Low-skilled" classes according to their average wage.

[16]It is a tool developed by Insee (`http://www.progedo-adisp.fr/apf_metric_en.php`)s, that uses data on the road network (`https://geoservices.ign.fr/bdtopo`) in order to commute travel time by car. The *commuting time* used in this paper is the time in minutes spent traveling using an individual car, off peak hours.

[17]An employer is an establishment.

[18]Note the long tail of commuting time for high-skilled workers: managers and engineers are indeed more likely to commute between metropolitan areas.

Figure 2 compares the evolution of commuting time, by broad occupation, using Davis et al. (2020)'s taxonomy of broad occupations, aggregated from the French 2-digit PCS-ESE classification (see section A.1), across eight different metropolitan areas. It clearly shows that same hierarchy in terms of commuting time, between broad occupations: high-skilled workers do commute more than middle-skilled ones, who in turn commute more than low-skilled ones - even though the differences between high- and middle-skilled workers can be small, as in Toulouse (AAU 005), Bordeaux (AAU 006) or Nantes (AAU 008).

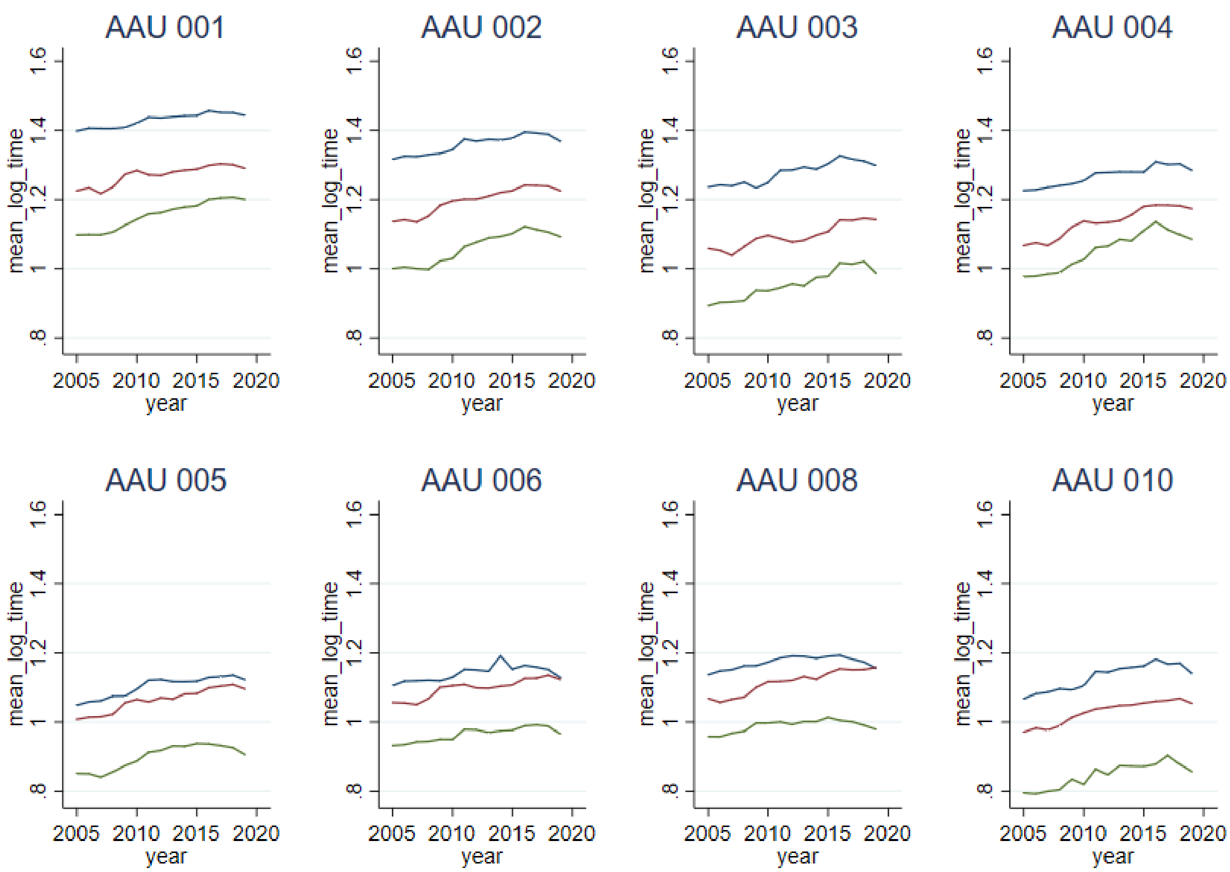


Figure 2: trends in log-commuting time by broad occupation, HS (blue)/ MS (maroon) / LS (green), in different metropolitan areas (001: Paris, 002: Lyon, 003: Marseille, 004: Lille, 005: Toulouse, 006: Bordeaux, 008: Nantes, 010: Strasbourg)

2.3 Labor demand factors account for a smaller fraction of wage and commuting time variance for the high-skilled

In order to capture, within a given metropolitan area, to what extent higher-wage earners of a given occupational group tend to commute more, we propose a simple metric: the partial correlation between the *unobserved wage component* and *unobserved commuting time component*, as estimated using a standard AKM regression (see next section 3).

We study this metric separately by broad occupational group, as managers and engineers ("high-skilled") exhibit distinct mobility patterns from both middle- and low-skilled ones, as Figure A.9 shows: the share of movers, that is, of workers with more than one employer over a 5-year period is 24% for the High-skilled, as compared with 32% and 34% for the Middle- and Low-skilled, respectively. Note that the raw fraction of movers can hardly be qualitatively comparable across broad occupational groups, as the data does not document whether such job moves are voluntary or endured - whether they are chosen by the worker or by the employer. A more plausible interpretation of such a difference between High-Skilled workers and Middle- and Low-Skilled ones is that the former have a greater bargaining power, so can benefit more from alternative offers, while staying in the same job (Cahuc et al. (2006)).

In Appendix A.3, we indeed show, in line with previous evidence,[19] that there is a clear hierarchy across occupations, in terms of the dispersion of wage and commuting time that is accounted for by labor demand factors – here defined, in a standard AKM framework, as *Employer Fixed Effects (EFEs)* and the *sorting* between workers and employers:[20] overall, these factors account for 15% of daily wage (7.1% of *commuting time*) dispersion for the High-Skilled, 24.7% (18.3%) for the Middle-Skilled, and 39.4% (27.2%) for the Low-Skilled. In other words, worker and employer fixed effects on *wage* and *commuting time* should be compared within a broad occupational category.

[19] See Cahuc et al. (2006).

[20] Which is defined as the normalized covariance between Worker Fixed Effects and Employer Fixed Effects.

3 AKM estimations at the a metropolitan area level

In this section, we describe how we measure unobserved heterogeneity, at the metropolitan area level. In a given a metropolitan area $\mathcal{P}$, we consider all employment matches where the employer (establishment) is located in $\mathcal{P}$, and all workers working for that establishment. Let us define the set of establishments located in $\mathcal{P}$ as $J(\mathcal{P})$. Following standard practice in the literature on two-way fixed effect estimations (see Bonhomme et al. (2023)), we restrict our estimation on the *largest connected set* of establishments, denoted as $J^c(\mathcal{P})$, where connections correspond to workers' job-to-job transitions.

3.1 The model

Formally, we consider any worker i who, over the 2010-2019 period, has been working at an establishment i located in $\mathcal{P}$, that is, all i such that there exists a year t so that $j(i,t) \in J^c(\mathcal{P})$, where $j(i,t)$ denotes worker i's unique employer j in year t.[21] Let us define this set of workers $\mathcal{I}^c(\mathcal{P})$. For all workers $i \in \mathcal{I}^c(\mathcal{P})$, we estimate *separately* worker-level fixed effects for the log-*daily wage rate* ω, on the one hand, and the log-*commuting time* τ, on the other, as dependent variables, in a standard linear additive AKM model with two-sided heterogeneity:

$$\begin{cases} \omega^\star_{it} &= \alpha^\omega_i + \psi^\omega_{j(i,t)} + X'_{ij(i,t)t}\beta + \varepsilon^\omega_{ij(i,t)t} \\ \tau^\star_{it} &= \alpha^\tau_i + \psi^\tau_{j(i,t)} + X'_{ij(i,t)t}\beta + \varepsilon^\tau_{ij(i,t)t} \end{cases} \tag{1}$$

where match control variables in X_{ijt} include:

- a cubic polynomial in the worker's potential market experience,[22]

[21] One observation in the data corresponds to a worker × year, in the sense that a worker is assumed to have only one employer in a given year t. This simplifying assumption is widely used in the AKM literature (see Bonhomme et al. (2023)).

[22] We are following Card et al. (2018) in this respect, even though we find it more relevant in the French case (as opposed to the case of Portugal in Card et al. (2018)) to restrict the age profile to be flat at 60 years old (as opposed to 40). See for that purpose the age profiles, depending on 3 different restrictions made in A.4, in the case of one metropolitan area.

– her gender, broad occupational category (high / middle / low-skilled)

– her current employer's 2-digit industry

– a year dummy variable, to account for the business cycle

Following the literature on AKM estimations (Bonhomme et al. (2023)), we cluster residuals ε^{ω} and ε^{τ} at the employer level.

We use the counterfactual notation of dependent variables as introduced in Bonhomme (2020): $(\omega^{\star}_{it}, \tau^{\star}_{it})$ denote *potential outcomes*, that is, potential daily wage and commuting time for a match formed by worker i and employer j at year t, regardless of whether such a match has actually taken place in the data. Let us denote by (ω_{it}, τ_{it}) the *observed outcomes*, which are linked to the potential ones through a dummy match variable D_{ijt}, which takes the value of 1 if a match between worker i and employer j at year t is present in the data: $(\omega_{it}, \tau_{it}) = (D_{ijt} \times \omega^{\star}_{it}, D_{ijt} \times \tau^{\star}_{it})$

We assume standards assumption when estimating AKM regressions 1, namely, *exogenous mobility*, in the form of a conditional orthogonality condition:

$$\begin{cases} \mathbb{E}\left[\varepsilon^{\omega}_{ijt} \mid D, X, \alpha^{\omega}, \psi^{\omega}\right] & = & 0 \\ \mathbb{E}\left[\varepsilon^{\tau}_{ijt} \mid D, X, \alpha^{\tau}, \psi^{\tau}\right] & = & 0 \end{cases} \tag{2}$$

This exogeneity assumption allows us to estimate Worker Fixed Effects $\{\alpha_i\}$ and Employer Fixed Effects $\{\psi_j\}$ using a standard OLS estimation. Let us denote by $\widehat{\alpha^{\omega}_{\mathcal{P}}}$ and $\widehat{\alpha^{\tau}_{\mathcal{P}}}$ the resulting estimates of Worker Fixed Effects (WFEs) on *daily wage* and *commuting time*, respectively, in a given metropolitan area $\mathcal{P}$, and by $\widehat{\psi^{\omega}_{\mathcal{P}}}$ and $\widehat{\psi^{\tau}_{\mathcal{P}}}$ the corresponding Employer Fixed Effects (EFEs).

The *exogenous mobility* assumption means that one takes as given the observed collection of matches formed, that is, the bipartite network structure between workers and firms,[23] and that, conditional on such a structure, the assignment of workers to firms is exogenous. In other words, this empirical model makes no use of a job ladder structure, as there is no explicit distinction between the observed distribution of matches, on the one hand, and a latent distribution of job offers, on the other, from which workers sample job opportunities and make optimal transitions accordingly.

Note that the exogenous mobility assumption does not mean that workers do not make optimal choices, when alternative job opportunities arise. To the contrary, it seems fair to posit that the observed distribution of matches, which we denote by G, results from workers' and employers' (constrained) optimal decisions, which are not directly observable. With this in mind, it becomes natural to interpret the estimated AKM fixed effects (in equation 1) through the lens of a job ladder model, which posits a latent job offer distribution, denoted by F, from which workers sample alternative job opportunities, resulting in a steady state distribution of observed matches, denoted by G.

3.2 Unobserved heterogeneity in the observed job and latent job offer distributions

Through the lens of a standard job-ladder model, workers sample job offers from a latent job offer distribution F, at a fixed job arrival Poisson rate, and, upon the reception of a new job offer drawn from a latent job offer distribution, chooses the privately optimal option, between either staying at her current job, or moving to the new job. The steady-state aggregation of all privately-optimal worker decisions generates an invariant, endogenous distribution, which corresponds to the observed empirical job distribution. Following the literature,[24] let us denote by G the empirical observed job distribution, and by F the latent job offer distribution.

[23] More specifically, the *connected component* of the bipartite network.

[24] See Manning (2003).

3.2.1 The observed job distribution G

Let us denote by $\widehat{\omega_{ij}}$ the unobserved, time-invariant component of the *observed* match (i, j)'s *daily wage*, and by $\widehat{\tau_{ij}}$ its unobserved, time-invariant component of the corresponding *commuting time*. Concerning workers in broad occupational group $\mathcal{O}$, in metropolitan area $\mathcal{P}$, these are defined by:

$$\begin{cases} \widehat{\omega_{ij}} & = \widehat{\alpha_i^{\omega}} + \widehat{\psi_j^{\omega}} \\ \widehat{\tau_{ij}} & = \widehat{\alpha_i^{\tau}} + \widehat{\psi_j^{\tau}} \end{cases} \qquad \text{where } o(i,t) = \mathcal{O} \text{ and } p\left(j(i,t)\right) = \mathcal{P} \tag{3}$$

We can then define G as the bivariate distribution of all observed matches' unobserved component of the *daily wage* and the *commuting time*, $\left(\widehat{\omega_{ij}}, \widehat{\tau_{ij}}\right)$.

3.2.2 The latent job offer distribution F

In a given metropolitan area $\mathcal{P}$, considering workers in a given broad occupational group $\mathcal{O}$, we allocate them *at random*, by *independently* bootstrapping observations of workers' fixed effects, on the one hand, and employers' fixed effects, on the other.[25]. We use the resulting matches to reconstruct the latent, unobserved job offer distribution $\mathcal{F}$. Using the counterfactual notation:

$$\begin{cases} \widehat{\omega_{ij}^{\star}} & = \widehat{\alpha_i^{\star,\omega}} + \widehat{\psi_j^{\star,\omega}} \\ \widehat{\tau_{ij}^{\star}} & = \widehat{\alpha_i^{\star,\tau}} + \widehat{\psi_j^{\star,\tau}} \end{cases} \qquad \text{where } o(i,t) = \mathcal{O} \text{ and } p\left(j(i,t)\right) = \mathcal{P} \tag{4}$$

Then we can define $\mathcal{F}$ as the bivariate distribution of all resulting counterfactual matches' unobserved component of the *daily wage* and the *commuting time*, $\left(\widehat{\omega_{ij}^{\star}}, \widehat{\tau_{ij}^{\star}}\right)$.

This definition makes use of the linear additive structure of the unobserved heterogeneity in the estimation model in 1: the unobserved, time-invariant, component of a

[25]Note that we bootstrap observations at the worker-time (i,t) level, not at the worker- or employer-level, following the literature (Bonhomme et al. (2023)).

match is the sum of the corresponding worker's and employer's time-invariant components, in logs. In other words, the cross elasticities are null: a 1% larger worker (*employer*) fixed effect will lead to a 1% larger match-level unobserved component, regardless of its employer's (*worker's*) identity. This simplifying assumption has been shown to be consistent with the data (Wong (2023)).

3.3 Wage-commuting trade-off at the metropolitan level: two measures

3.3.1 The Commute-Wage Gradient (CWG)

The Commute-Wage Gradient (CWG) is the empirical measure corresponding to the slope of a representative worker's indifference curve, in the log-Commute/Wage plane. It corresponds to the hypothetical commute/wage trade-off that would be faced by a marginal incoming worker of a given broad occupation, in a given metropolitan area.

Assuming linearity in the empirical commute/wage unobserved heterogeneity schedule, the Commute-Wage Gradient (CWG) in a given metropolitan area $\mathcal{P}$ and for a given broad occupation $\mathcal{O}$ (high-/middle-/low-skill), is estimated as the coefficient $\beta_{CWG}^{\mathcal{P},\mathcal{O}}$ in the simple OLS:

$$\widehat{\tau_{ij}} = \beta_{CWG}^{\mathcal{P},\mathcal{O}} \widehat{\omega_{ij}} + \beta_0 + \varepsilon_{ij} \tag{5}$$

for all (i,t) such that $o(i,t) = \mathcal{O}$ and $p(j(i,t)) = \mathcal{P}$.[26] The $\left(\widehat{\omega_{ij}}, \widehat{\tau_{ij}}\right)$ come from equation 3.

3.3.2 The Monopsony Power Measure (MPM)

In the spirit of the job ladder literature, we define the Monopsony Power Measure (MPM) as the inverse of the *relative distance* between the observed Commute-Wage Gradient (CWG), and the counterfactual one, which we label CWG*. We define the latter

[26] Note that we abstract from non-linearities at both ends of the distribution by restricting the estimation of $\beta_{CWG}^{\mathcal{P},\mathcal{O}}$ to the middling 60% of the ω distribution.

in an analogous way as the observed CWG: the counterfactual Commute-Wage Gradient (CWG*) in a given metropolitan area $\mathcal{P}$ and for a given broad occupation $\mathcal{O}$ (high-/middle-/low-skilled), is estimated as the coefficient $\beta^{\mathcal{P},\mathcal{O}}_{CWG^\star}$ in the simple OLS:

$$\widehat{\tau^\star_{ij}} = \beta^{\mathcal{P},\mathcal{O}}_{CWG^\star}\, \widehat{\omega^\star_{ij}} + \beta^\star_0 + \varepsilon^\star_{ij} \tag{6}$$

for all (i,j,t) such that $o(i,j,t) = \mathcal{O}$ and $p(j,t) = \mathcal{P}$. The $\left(\widehat{\omega^\star_{ij}}, \widehat{\tau^\star_{ij}}\right)$ are defined in equation 4.

The Monopsony Power Measure (MPM)$^{\mathcal{P},\mathcal{O}}$, in a given metropolitan area $\mathcal{P}$ and for a given broad occupation group $\mathcal{O}$, is defined as the inverse of the *relative distance* between the two Commute-Wage Gradients defined above. More specifically, it takes the *latent distribution of job offers* as the baseline, taking the inverse of relative deviation (in percentages) of CWG with respect to CWG*.[27]

$$\text{MPM}^{\mathcal{P},\mathcal{O}} = \left(\frac{\beta^{\mathcal{P},\mathcal{O}}_{CWG}}{\beta^{\mathcal{P},\mathcal{O}}_{CWG^\star}} - 1\right)^{-1} \tag{7}$$

3.4 Results

Preliminary result: the empirical MPM is consistent with job-ladder models

In practically all metropolitan areas, and for all three broad occupational groups, the counterfactual Commute-Wage Gradient is lower than the observed one:

$$\beta^{\mathcal{P},\mathcal{O}}_{CWG^\star} < \beta^{\mathcal{P},\mathcal{O}}_{CWG} \tag{8}$$

This empirical finding ensures that our MPM measure is consistent with a job ladder model, whereby workers sample job offers from a latent distribution, and choose the optimal option between staying and switching jobs. Indeed, in the basic job ladder model

[27] Indeed, since the lower the distance of CWG with respect to CWG*, the greater the monosponistic power, we consider the inverse of the relative distance.

with only one variable, the wage, the observed job distribution is derived from the latent job offer distribution, through a left truncation of the latter, as workers seek higher-wage opportunities.[28] With commuting time as an additional variable in F and G, the marginal distribution of commuting time for the observed job distribution will result from the marginal job offer distribution through a right truncation of the latter, as workers seek shorter-commute opportunities. Hence, optimal worker decisions concerning both the wage and the commuting time imply that the Commute-Wage Gradient in the observed G distribution, CWG, should be larger than the Commute-Wage Gradient in the latent F distribution, CWG*.

Figure 3 shows an example of a measure of CWG and CWG*, in the case of high-skilled workers in Nantes.

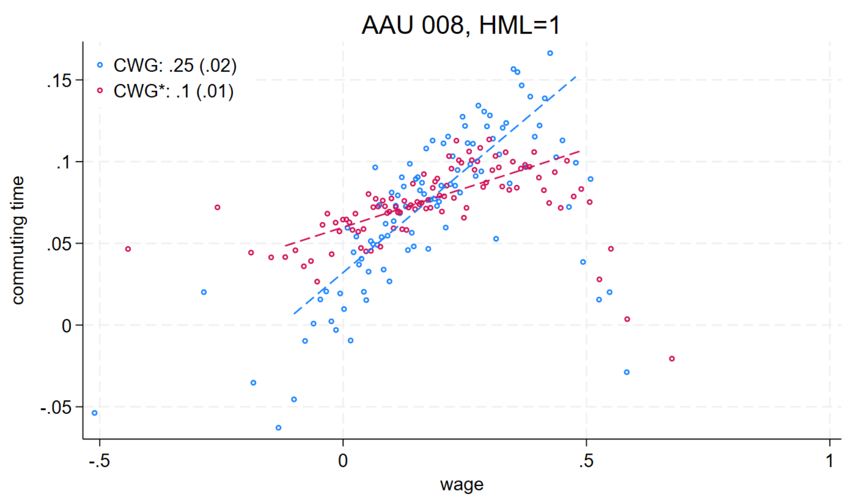
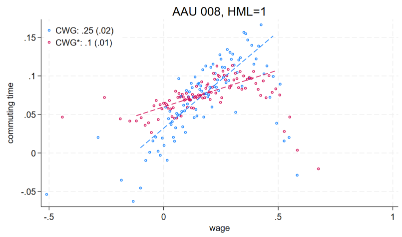


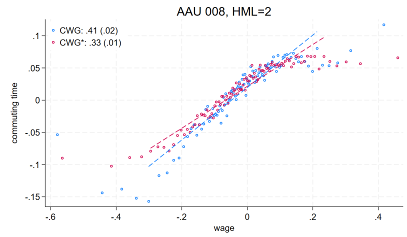
Figure 3: Binned scatterplots of the *observed distribution of matches* ($(\widehat{\omega_{ij}}, \widehat{\tau_{ij}})$, in blue), and of the *latent distribution of matches* ($(\widehat{\omega^\star_{ij}}, \widehat{\tau^\star_{ij}})$ in red). CWG is the slope of a fitting line for the blue dots, while CWG* is the slope of a fitting line for the red dots, estimated on the 5^{th} to 95^{th} percentile interval along the x-axis. The metropolitan area is Nantes.

Before presenting the results across all 13 metropolitan areas, let us consider the Commute-Wage Gradients for all three occupational groups, in Nantes, as shown in Figure 4.

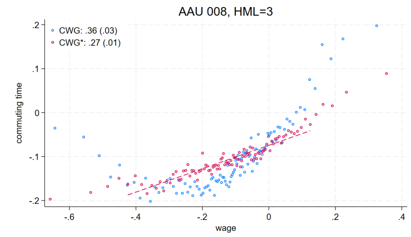
[28]See equation (2.18) in Manning (2003).



(a) High-skilled (1)



(b) Middle-skilled (2)



(c) Low-skilled (3)

Figure 4: Binned scatterplot of residual observed *wage* distribution ($\widehat{\omega}$, x-axis), and on *commuting time* ($\widehat{\tau}$, y-axis), by broad occupational group (high-skilled/middle-skilled/low-skilled), in Nantes

The case of Nantes metropolitan area illustrates the general patterns which are to be found across practically all 13 metropolitan areas considered. First, the Commute-Wage Gradient (CWG) is highest for middle-skilled workers (Figure 4b), as compared to both high-skilled (Figure 4a) and low-skilled workers (Figure 4c), so that the marginal middle-skilled entrant in Nantes faces a more constrained commute-wage trade-off: were she to seek higher wages, she would have to commute more, than a marginal high-skilled or low-skilled entrant. Second, the relative distance between CWG* and CWG is smallest for the Middle-Skilled, so that the Monopsony Power Measure (MPM) is highest for these workers. Indeed, a small relative distance means that workers hardly benefit from utility-improving job switch opportunities, that is, from on-the-job search. By contrast, high-skilled workers do benefit from alternative job offers, as the relative distance between CWG* and CWG is higher. Last, low-skilled workers exhibit a U-shaped pattern in the observed commute-wage schedule (blue dots in Figure 4c). The downward sloping part to the left of the plot is hardly consistent with a job ladder model, as it means that lower wage workers commute more than median-wage ones.

Result 1: Middle-skilled workers both face a tighter commute-wage trade-off, and greater monopsony power

Figure 5 offers a synthetic picture, across 13 metropolitan areas considered here, of the Commuting-Wage Gradient (CWG) for the three broad occupational groups. The

CWG measured for the middle-skilled workers (in red) is higher than the one measured for the high-skilled (in blue) in most metropolitan areas, with the notable exception of Nice, as well as Toulon and Nancy where they are statistically close. The difference with the low-skilled workers' CWG is less marked, with the middle-skilled estimates being larger in 7 out of 13 metropolitan areas. Note, though, that the CWG for the low-skilled should be taken with caution, due to the non-monotonicities that their Commute-Wage schedule exhibits. Second, middle-skilled workers face stronger monopsony power than high-skilled ones, as Figure 6 shows.

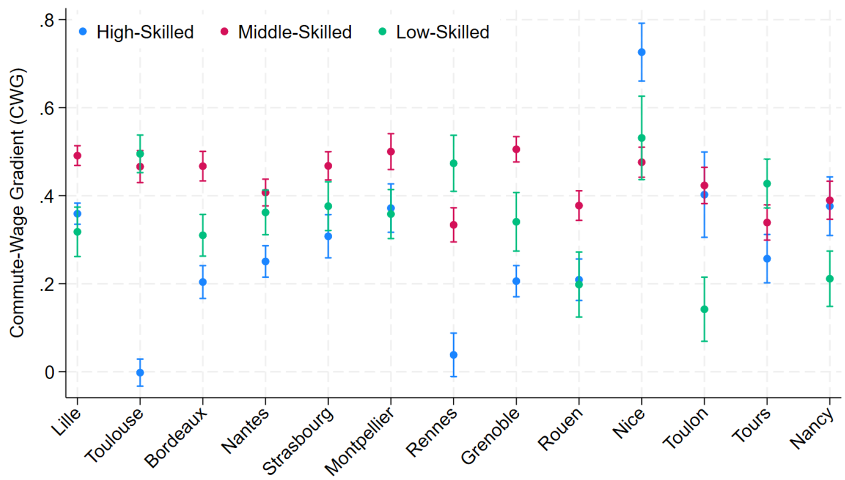


Figure 5: Commute-Wage Gradient (CWG) estimates, with 95% confidence intervals, observed job distribution, across 13 metropolitan areas

Result 2: High-skilled workers are less constrained in their commute-wage trade-off, and face less monospony power, than middle- and low-skilled workers

Figure 6 shows that high-skilled workers face less monopsony power than the middle-skilled ones in practically all metropolitan areas, with the exceptions of Grenoble and Bordeaux.

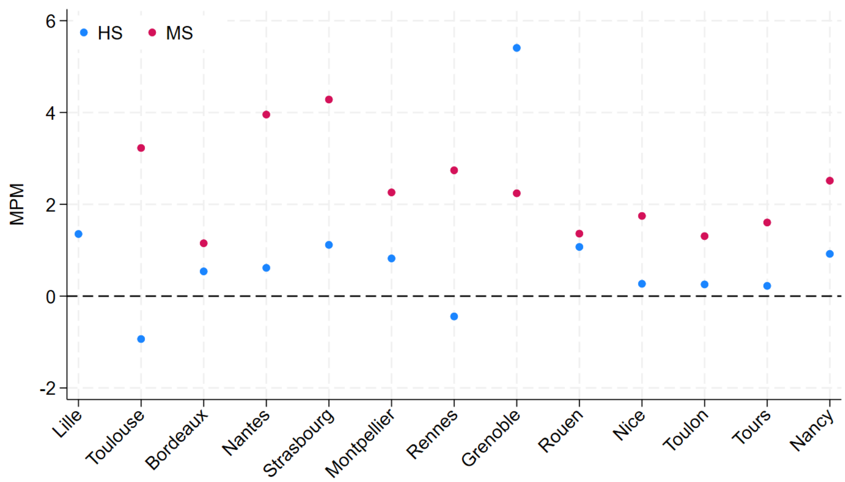


Figure 6: Monopsony Power Measure (MPM) as defined in equation 7, for the high- (blue) and middle-skilled workers (red).

Result 4: Low-skilled workers' commute-wage schedule exhibits strong strong non-monotonicities

Low-skilled workers overall face a qualitatively different Commute-Wage trade-off, in all 13 metropolitan areas considered, as is evident from Figures A.2, A.3, A.4 and A.5. More specifically, it exhibits a U-shaped pattern, with lower wage earners having to commute more than median-wage ones, who in turn commute less than higher wage earners. At this point of investigation, we can only propose a tentative interpretation for such a pattern, namely: low-skilled workers tend to reside the surrounding *exurbs*, where wages are smaller than in the denser, principal cities of a metropolitan area, so that workers who commute the least earn median wages, living and residing in the same town, while commuters either finds jobs in the surrounding area, at lower wages (to the left on the x-axis), or in the principal city, where wages are higher (to the right).

Conclusion

We show here how unobserved heterogeneity, measured using the canonical two-way fixed effect additive linear model, sheds new light on workers' commute-wage trade-off, depending on their broad occupational group. The two measures proposed here, following the logic of the dynamic monopsony literature, consistently show that middle-skilled workers both face an overall more constrained trade-off, than high- and low-skilled ones, and are more subject to employers' monospony power, as opposed to high-skilled workers, who fare better according to both measures across practically all metropolitan areas considered. Last, low-skilled workers stand in stark contrast to the other two occupational categories, with a commute-wage schedule displaying strong non-monotonicities, which calls for further investigation.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," Econometrica, 67, 251–333.

Babet, D., O. Godechot, and M. Palladino (2023): "In the Land of AKM: Explaining the Dynamics of Wage Inequality in France," .

Bonhomme, S. (2020): "Chapter 5 - Econometric analysis of bipartite networks," in The Econometric Analysis of Network Data, ed. by B. Graham and Áureo de Paula, Academic Press, 83–121.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2023): "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" Journal of Labor Economics, 41, 291–322.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," Econometrica, 74, 323–364.

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.

Davis, D. R., E. Mengus, and T. K. Michalski (2020): "Labor Market Polarization and the Great Divergence: Theory and Evidence," Working Paper 26955, National Bureau of Economic Research.

Farré, L., J. Jofre-Monseny, and J. Torrecillas (2022): "Commuting time and the gender gap in labor market participation," Journal of Economic Geography, 23, 847–870.

French, M. T., I. Popovici, and A. R. Timming (2020): "Analysing the effect of commuting time on earnings among young adults," Applied Economics, 52, 5282–5297.

Gimenez-Nadal, J. I., J. A. Molina, and J. Velilla (2018): "The commuting behavior of workers in the United States: Differences between the employed and the self-employed," Journal of Transport Geography, 66, 19–29.

Heblich, S., A. Trew, and Y. Zylberberg (2021): “East-Side Story: Historical Pollution and Persistent Neighborhood Sorting,” Journal of Political Economy, 129, 1508–1552.

Inoa, I. A., N. Picard, and A. de Palma (2015): “Effect of an Accessibility Measure in a Model for Choice of Residential Location, Workplace, and Type of Employment,” Mathematical Population Studies, 22, 4–36.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2020): “Gender Differences in Job Search: Trading off Commute against Wage*,” The Quarterly Journal of Economics, 136, 381–426.

Manning, A. (2003): Monopsony in motion: imperfect competition in labor markets, Princeton University Press.

Sang, S., M. O’Kelly, and M.-P. Kwan (2011): “Examining Commuting Patterns: Results from a Journey-to-work Model Disaggregated by Gender and Occupation,” Urban Studies, 48, 891–909.

Wong, H. C. (2023): “Understanding High-Wage Firms,” Working paper.

A Appendix

A.1 Classification of French 2-digit occupations into three broad occupational categories, from Davis et al. (2020)

Table 1 – Sample statistics by 2 digit CS categories.

CS	Description	Employment Share percent		Average City Wage (in 2015 euros)		Routine	Offshorable
						rankings	
		1994	2015	1994	2015		
		high-paid occupations					
23	CEOs	1.0	0.9	42.81	59.20	16	17
37	**managers and professionals**	**6.2**	**10.2**	32.52	38.56	15	16
38	**engineers**	**5.1**	**9.0**	30.36	33.69	17	10
35	creative professionals	0.5	0.5	22.83	31.80	14	11
		middle-paid occupations					
48	**supervisors and foremen**	**4.1**	**2.7**	18.03	21.86	3	3
46	**mid-level professionals**	**12.3**	**7.6**	17.54	21.20	13	6
47	**technicians**	**5.7**	**6.3**	17.15	20.60	11	7
43	mid-level health professionals	0.8	1.5	15.05	18.05	10	13
62	**skilled industrial workers**	**14.1**	**9.3**	13.52	17.99	4	2
54	**office workers**	**11.8**	**11.2**	13.17	16.98	1	4
65	**transport and logistics personnel**	**2.9**	**3.0**	11.96	16.00	5	5
63	**skilled manual workers**	**8.0**	**8.3**	11.90	15.50	7	8
64	**drivers**	**5.0**	**5.5**	11.50	14.46	18	18
67	**unskilled industrial workers**	**10.9**	**5.7**	11.02	14.72	2	1
		low-paid occupations					
53	security workers	0.7	1.4	10.60	14.60	9	12
55	**sales-related occupations**	**5.4**	**8.3**	10.44	13.74	6	15
56	personal service workers	2.2	4.8	9.97	12.63	12	14
68	**unskilled manual workers**	**3.3**	**3.8**	9.11	13.28	8	9

Figure A.1: Occupation classification from Davis et al. (2020)

A.2 Comute-Wage patterns across 3 broad occupational groups, in 13 metropolitan areas (HML = 1 (high-skilled) / 2 (middle-skilled) / 3 (low-skilled))

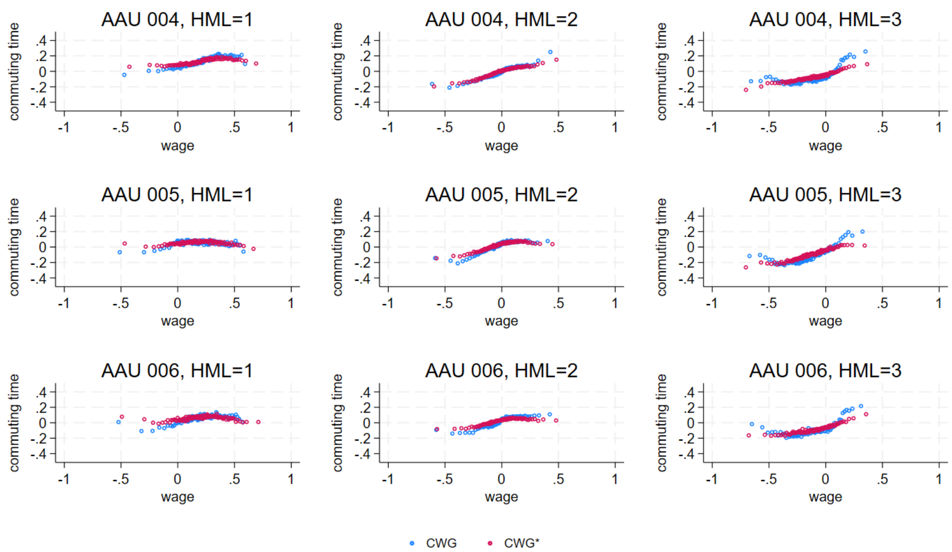


Figure A.2: Commute-Wage schedule in the observed distribution (blue) and in the latent job offer distribution (red), Lille (top panel), Toulouse (middle) and Bordeaux (bottom)

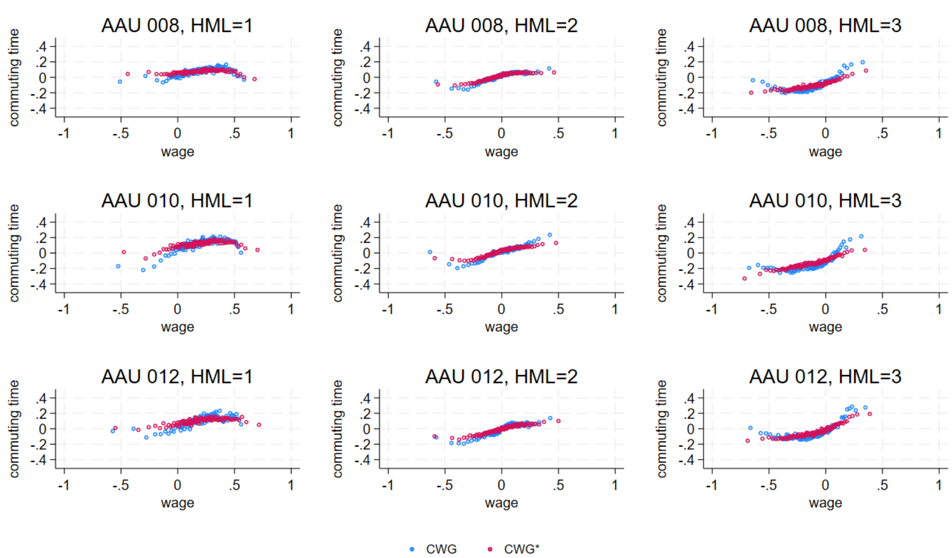


Figure A.3: Commute-Wage schedule in the observed distribution (blue) and in the latent job offer distribution (red), Nantes (top panel), Strasbourg (middle) and Montpellier (bottom)

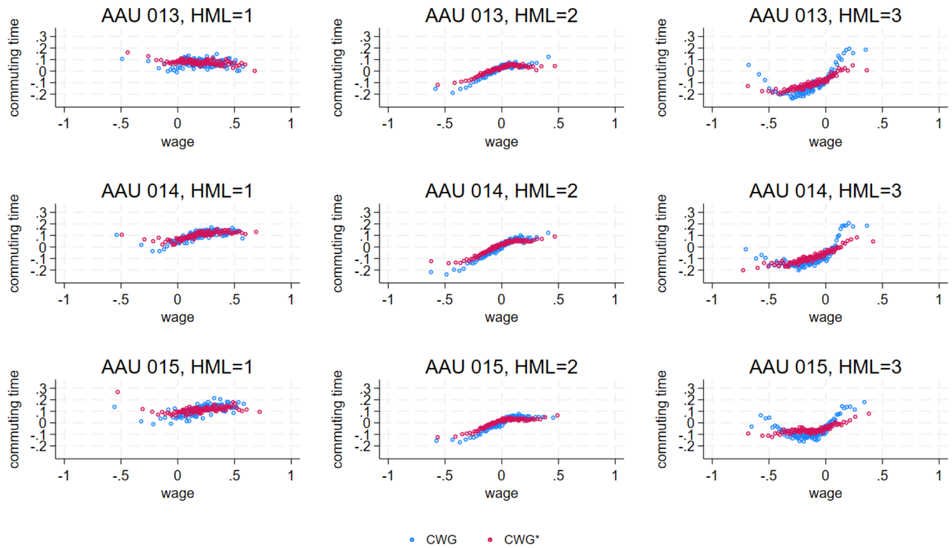


Figure A.4: Commute-Wage schedule in the observed distribution (blue) and in the latent job offer distribution (red), Rennes (top panel), Grenoble (middle) and Rouen (bottom)

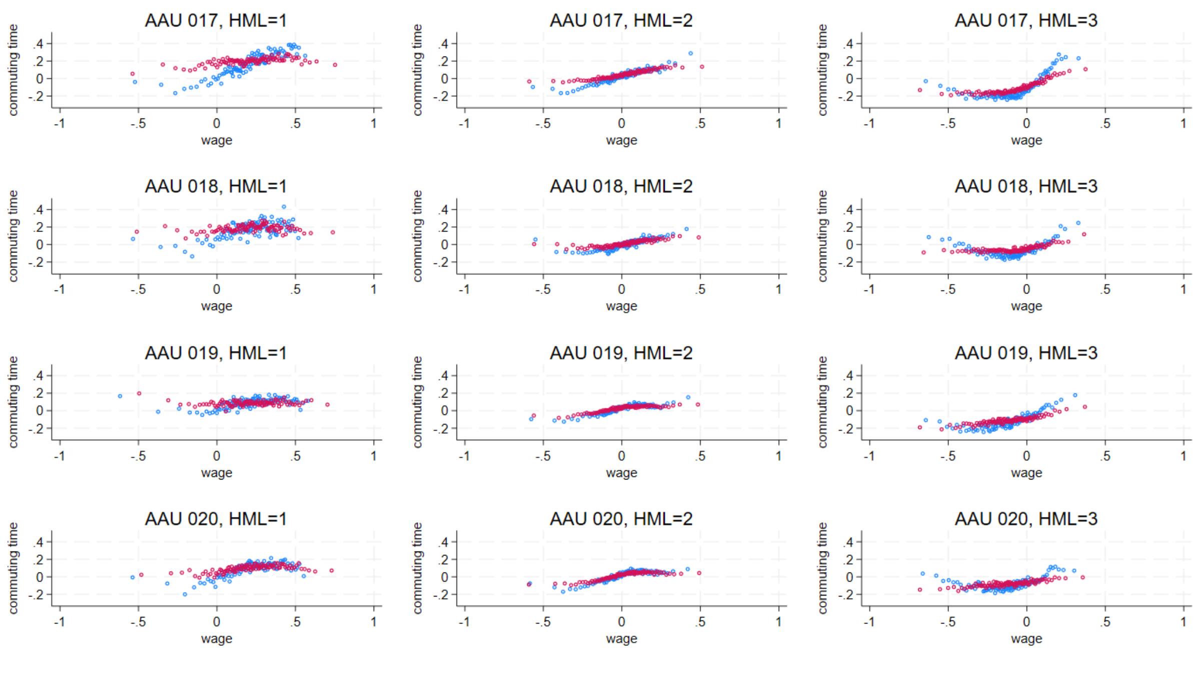


Figure A.5: Commute-Wage schedule in the observed distribution (blue) and in the latent job offer distribution (red), Nice (first panel), Toulon (second), Tours (third) and Nancy (fourth)

A.3 AKM variance decomposition: contributions of labor demand factors to the dispersion of wage and commuting time

In this section we assess the respective contributions of employer fixed effects (EFEs) and sorting to wage- and commuting time dispersion at the metropolitan area level (among the 16 largest ones in France). The estimates presented here results from a *correlated random effects* estimator as proposed by Bonhomme et al. (2023), where employers are in a first stage clustered into a fixed number of classes, using a K-means algorithm. Discretizing unobserved heterogeneity on employers' side indeed has the attractive feature that it mechanically increases workers' moves between classes of employers, so that estimates are robust to the well-known *limited mobility bias* and are therefore comparable across metropolitan areas and occupations which differ in worker mobility rates.

Indeed, Figure A.9 shows that, overall, Middle- and Low-skilled workers have a larger share of movers,[29] than High-skilled ones. Note also that worker mobility, as measured by the share of movers among workers, is more sensitive to the size of the pool of workers (x-axis) for the High-skilled, than for the Middle- and Low-skilled.

[29]Defined as workers who have more than one employer over the period 2015-2019.

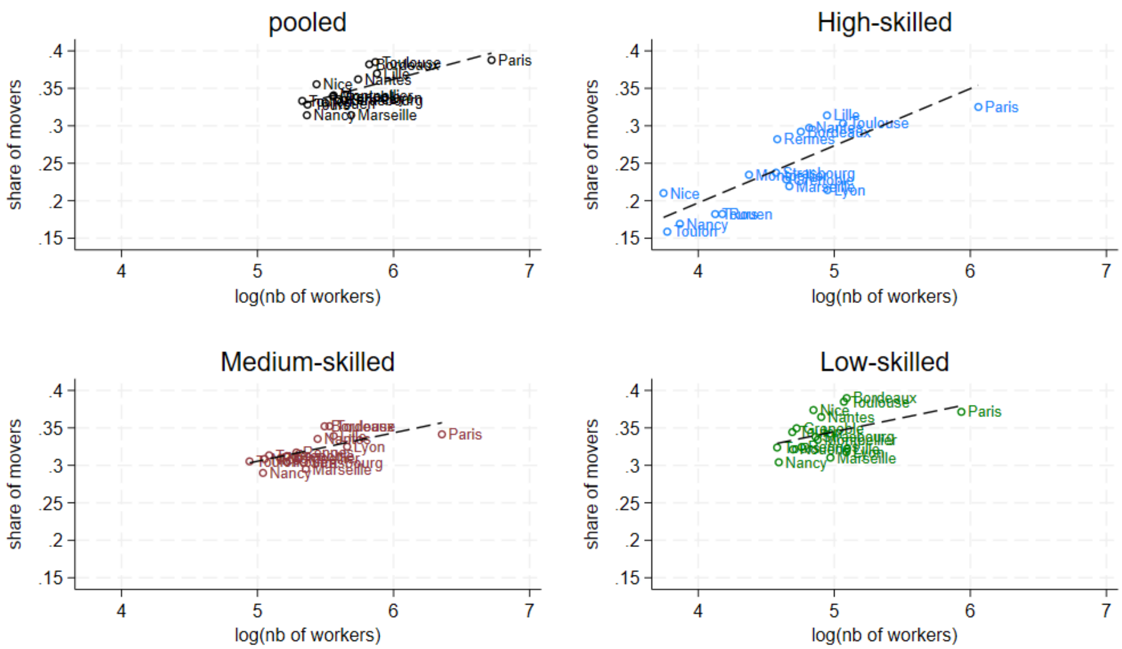


Figure A.6: share of movers and number of workers by occupation and metropolitan area, over the period 2015-2019

average ("0") and by broad occupation ("1"=H) at the metropolitan area level

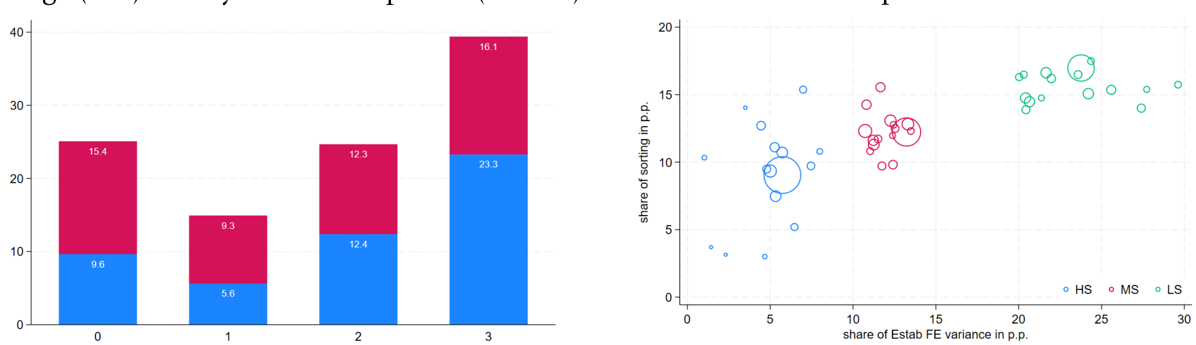


Figure A.7: contributions of *sorting* and *EFEs* to the dispersion of workers' *daily wage rate*, 2015-2019

Figure A.7 shows that the contribution of Employers Fixed Effects (EFEs) to the dispersion of workers' *daily wage rate* is larger for Low-Skilled workers (23.3%), than for Middle-Skilled workers (12.4%), than for High-Skilled workers (5.6%). The contribution of *sorting* roughly follows the same hierarchy, though the differences across occupational groups are less pronounced (16.1%, 12.3% and 9.3%, respectively). The estimated difference across occupations is in line with previous research (see Cahuc et al. (2006)).

Importantly, the right panel shows that this hierachy across occupation is robust at the metropolitan area level: within a given area, sorting and employer effect will contribution more to the dispersion of Low-Skilled wage rates, than to that of Middle-Skilled wage rates, than to that of High-Skilled wage rates.

average ("0") and by broad occupation ("1"=H) at the metropolitan area level

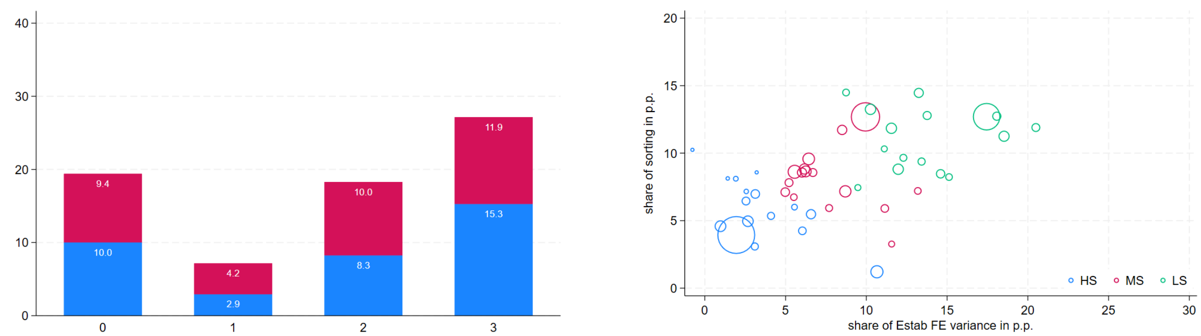


Figure A.8: contributions of *sorting* and *EFEs* to the dispersion of workers' *commuting time*, 2015-2019

Figure A.8 shows that a similar hierarchy across occupational groups also holds in the case of workers' dispersion in *commuting time*: both *sorting* and *Employer Fixed Effects* (EFEs) contribute less to the dispersion of commuting time for High-Skilled workers, than for Middle-Skilled ones, than for Low-Skilled ones.

Note, though, that labor demand factors (*sorting* and *EFEs*) contribute less overall to the dispersion of workers' commuting time, than to the dispersion of workers' wage: overall, sorting and EFEs account for 25% of wage rate dispersion, as opposed to 19.5% of commuting time dispersion. The difference between wage and commuting is most striking among High-Skilled workers: these two factors account for 15% of wage rate dispersion, as opposed to 7.1% for commuting time.

A.4 Age profile, as implied by equation (1), in Grenoble

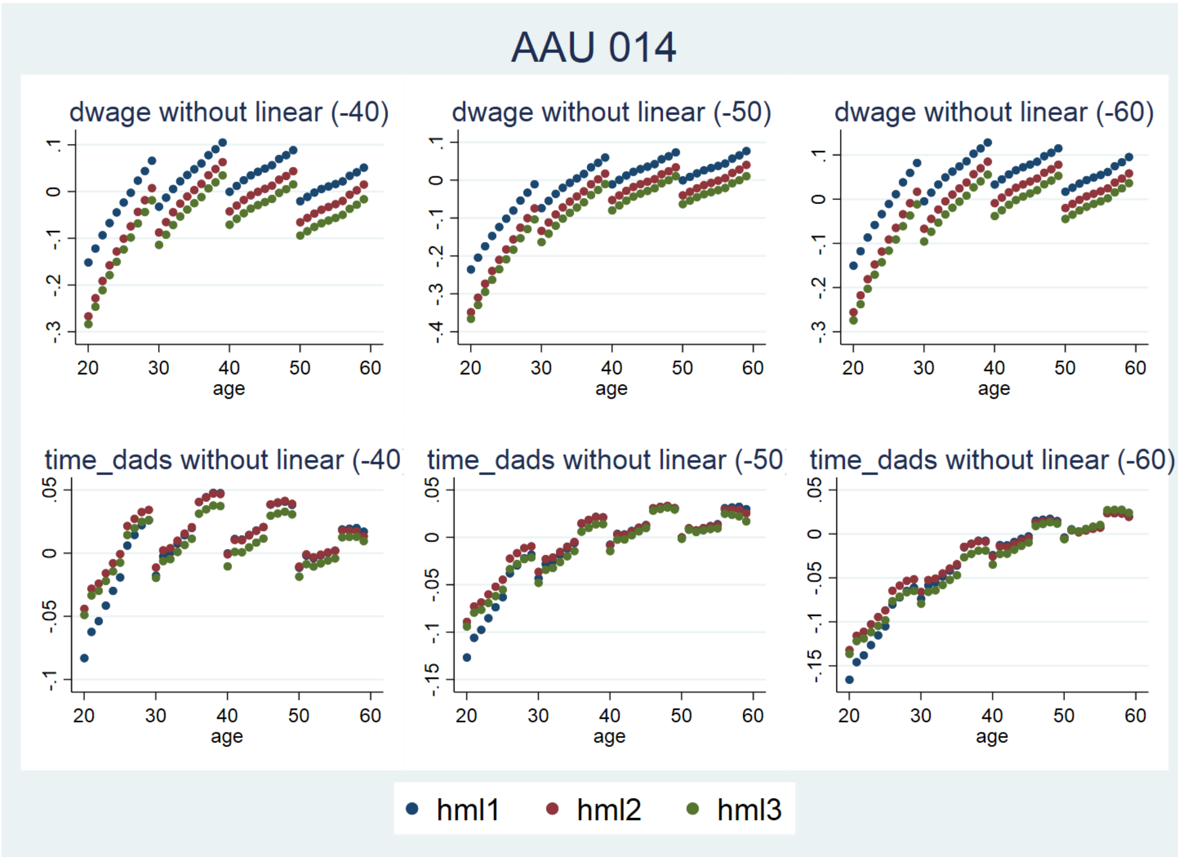


Figure A.9: implied age profile, by 10-year cohort, Grenoble