High Wage Workers and High Amenity Firms

Romain Angotti*

May 17, 2024

Abstract

Amenities are unobservable characteristics of jobs. The literature generally finds that including them augments total compensation inequality in the economy. I build on the framework from Sorkin (2018) to identify amenities at the firm level. My model includes worker heterogeneity which generates differential mobility rates in the economy. I estimate it using a structural identification from flows of workers across firms. Using data on the French labour market, I show that including amenities leads to more dispersion in utility across firms than suggested with wage data only. I find that including differential mobility rates across workers exacerbates this dispersion because good workers are more likely to move up the firms-ladder than less productive workers. Sorting on unobserved job characteristics is even stronger than on firms wage premia. In addition, I find that the verticality of the firms-utility-ladder is salient only for low income workers. On the contrary, for higher income workers, wages and amenities are correlated negatively which suggests compensating differentials across firms.

1 Introduction

Identifying the non-wage component of a worker's compensation is hard with only wage data. However, these unobservables, such as the quality of the working environment, the professionalism of the management, the flexibility of the work schedule or even one's autonomy can add up to represent a large part of what constitutes the total compensation for a job. It happens that some of these characteristics are correlated negatively with the wage, for example when workers are compensated for fatality risk (Lavetti and Schmutte 2018). On the contrary, some are not priced in the wage and therefore add up positively to the worker's total compensation (Maestas et al. 2023). Ultimately the question of whether a certain non-wage characteristic of a job adds up or subtracts to the wage is an empirical question. Both have been documented and are theoretically possible (Sorkin 2018).

The role of firms in the dispersion of wages in the economy has long been studied. Since Abowd et al. (1999) (henceforth AKM), economists have tried to empirically measure the

*This work constitutes my master thesis for the Sciences Po research master in economics. I thank my advisor Jean-Marc Robin and jury Yann Algan for their helpful guidance. I am extremely greatfull to Philippe Aghion and Antonin Bergeaud for their help and for getting me involved in a related project that inspired this work. Most of this master thesis was prepared from the Innovation Lab at Collège de France where I received help and support from all my colleagues, in particular Francis Kramarz, Corentin Laffitte and Anaïs Moreau. I also want to thank all my friends and professors at Sciences Po for their help and advice, in particular Nicolas Dussaux and Constantin Gleyze. I address additional thanks to Tobias Lehmann and Pierre-Philippe Combes for helpful discussions. Finally, I infinitely thank my family and friends for their unwavering non-economics support. This work uses data produced by the French national statistical agency which I accessed through the Centre Sécurisé d'Accès aux Données.

sources of wage inequality in the labour market. This literature finds that between firms differences are very thin once corrected for the fact that workers are more productive in higher-paying firms than in lower paying firms. The consequence of this sorting of workers is that firms do not particularly seem to be driving inequality as compared to differences between workers themselves.

However, disregarding job amenities and dis-amenities at the firm level might confound firm wage premia with Compensating Differentials (CDs) at the firm level. If a firm pays a premium because working there is awful, it does not contribute to inequality in total compensation, which is the relevant welfare measure in fine. Sorkin (2018) uses a reduced-form[1] choice model and finds that 70% of the firm wage premia reflect CDs. Hence, once one controls for the total compensation differences by correcting for the unobserved amenity, firm pay differences are even thinner than previously thought.

Since the correlation between firm amenities and firm wage premia is not exactly -1, differences in unobserved amenities add to inequality even though not through wage differentials. If amenities and firm wage premia correlate positively, workers face a firm-utility-ladder. Going up this ladder increases workers utility since it increases wages and amenities. The objective of this work is to document whether CDs or the Utility-Ladder (UL) dominate in the data when considering between firms differences. In particular, I build on Sorkin's model to measure total utility dispersion in the economy. In his vein, I assume that firms provide a common amenity to all workers within the firm.

The role of amenities in labour markets has recently received a revival of attention. In particular, Maestas et al. (2023) find in a survey that US workers' unobserved job characteristics increase the dispersion of total compensation in the economy by 10 log points. They include the amenity bundle to the wage by pricing each amenity according to the stated preferences elicited in their survey. Sockin (2021), by collecting a large dataset on amenities from the platform Glassdoor on which workers rate their employers, elicits a dispersion in total compensation 20% higher than for wages only across US firms. On the high end, Lehmann (2023), using a model closer to Sorkin's, finds that amenities add to inequality in Austria by around 60%.

In this work, I consider all non-wage characteristics that are attached to a job. Some jobs can give a sense of self-esteem or provide learning opportunities whereas some are exhausting, stressful or dangerous. Since many different dimensions could contribute to the amenity bundle, one unifying methodology is to measure the value of the all non-wage attributes of a job. This is the approach I use in this work. Previously, this heterogeneity has been thought in terms of inter-industry differentials (Krueger and Summers 1988) since they drive most of the differences in the occupational content of jobs.

This definition goes in the direction of several other works (Hall and Mueller 2018 or Lehmann 2023) but differs from the alternative definition of amenities as the idiosyncratic worker preference for a job[2]. This other literature considers that the attractiveness of a job comes from subjective preferences that are peculiar to each worker. For example, the distance between the workplace and one's home can deliver utility to workers without being a shared amenity that the firm provides to all workers. In general, this is thought of as a private rent of workers since the firm cannot observe it.

In reality, amenities most often fall in between the idiosyncratic and the shared categories. This ambiguity comes from the fact that workers have heterogeneous tastes for amenities and

[1] He compares expected utility V^e with firm fixed effect in the wage.
[2] notably Lamadon et al. (2022) or Berger et al. (2023)

sort in jobs that appeal most to their preferences. Maestas et al. (2023) indeed find that heterogeneity in preferences induces a significant sorting of workers depending on their tastes for different amenities.

More generally, identifying preferences in the labour market is challenging because of its frictional nature. This has made the task of economists to identify CDs very difficult. Bonhomme and Jolivet (2009) show how amenities play a significant role in worker choices but that frictions in the labour market would need to be extremely low for wages to exhibit CDs. The subsequent literature therefore puts more intense use on choice models to identify them. This is the case of Sullivan and To (2014a) and Taber and Vejlin (2020). Those models include several channels that make identification of amenities complex: human capital, comparative advantage and frictions. My framework is more in line with Sorkin (2018) whose model only features frictions and makes no use of wage information.

Sorkin (2018) uses a ranking algorithm that reconciles firm values and observed flows of workers across firms. His model is built on a revealed preferences argument according to which the direction of the in- and out-flows of workers from each firm indicates how this firm stands in the overall ranking of firms. This technique builds on the PageRank algorithm developed by Google (1998) for ranking webpages. He then carefully accounts for potential shocks that affect mobility exogenously and would otherwise confound the revealed preference of workers. I depart from him by not correcting for those potential biases using additional moments of the data but instead by estimating my model with the more restricted sample of tenured workers. Tenured workers are somewhat insulated from involuntary job changes, therefore their mobility decisions reveal more closely their preferences. Regarding the firm offers distribution, Sorkin identifies firm recruiting intensity by using hires from non-employment. Evidence from Faberman et al. (2022) show that employed and unemployed sample job offers from different distributions. Their findings contradict this identifying assumption so I prefer making the more agnostic assumption that firms meeting probabilities are uniform.

I add to his model by incorporating worker heterogeneity. The idea is that flows of workers are not all worth the same because all workers are not as likely to be poached. As in Lamadon et al. (2022), workers are heterogeneous in their individual efficiency. I allow firms to impose a markdown on wages and extract more profits from good types. This delivers a hiring probability that depends on the type of the workers. As a result, good workers are more mobile and more likely to climb the firms-ladder.

With this framework, I can speak to the entire utility dispersion in the economy, factoring-in the unobservable amenities, and recover granular firm-level estimates. I describe the full distribution of firm wage premia, firm amenities and firm average worker productivities. My model nests the homogeneous workers framework of Sorkin as a simpler case. I also recover the labour market parameters that govern mobility and explain the sorting I observe. Sorting in my model is not an equilibrium outcome but results from frictions and the differential mobility of workers.

I show how differential mobility affects the expected value workers have for firms. The utility dispersion is greater for low productivity workers because they are more likely to be stuck in a bad firm. I then show how to estimate my model using a matched employer-employee dataset. The first step is to perform an AKM regression to identify worker and firm fixed-effects in the wage. Then I use the flows of workers across firms and the average worker qualities to estimate my model. The estimation is iterative since my firm values depend on mobility parameters that can only be estimated with some guess for firm values.

I have to approximate my model by estimating it for the average type in each firm - which corresponds to half of the variance in types in the economy. Sorkin's method is supposedly faster than using a Maximum Likelihood Estimator as is done in most of the literature. This is because it only deals with the very sparse matrix of between-firms flows which is $K \times K$ for K firms even in the presence of millions of workers. However, by including non-linearities due to worker heterogeneity, the computational advantage is less clear.

I take my model to the French Ile-de-France DADS[3] data for the period 2010-2015. Since I restrict myself to the well-connected set of firms, my AKM estimates are fairly robust to the limited mobility bias[4]. I find most of the wage variation to be explained by worker-effects and a strong correlation of worker types with firm fixed effects. As Sorkin, I find CDs by identifying the portion of the firm effect that equalizes utility differences. Compensating workers at the firm level for differences in firm amenities decreases the firm effect by 47%, less than what Sorkin (2018) finds for the the US. However my estimate using his homogeneous workers model are fairly close.

Higher productivity workers are found to be more mobile than the average. I identify the 3 parameters of my model that relate to mobility and utility dispersion in the French economy. The first one governs overall frictions in the economy, the second how mobility varies with worker types and the third is the variance of the idiosyncratic preferences relative to the total differences in utility across firms. I find that mobility is very sensitive to worker types. This implies that utility in the economy is significantly more dispersed than suggested by a homogeneous workers model because bad workers cannot relocate as easily as good workers. In addition, I find that the idiosyncratic component of utility has a low variance compared to the size of the firms-utility-ladder. I measure this parameter by scaling utility differences measured relative to the scale of the taste shock to the observed differences in firm premia of movers that go through unemployment. This derives from the additional assumption that, when unemployed, workers put more emphasis on wages than amenities when considering job offers.

My results align with observable firm characteristics. I compare amenities to sectors and locations of firms and find that their distribution makes intuitive sense. A fine industry decomposition explains as much as 45% of the differences in amenity-provision whereas it leaves wide differences in wage premia across firms. I also check that the amenity is not entirely explained by differences in working time.

I provide several decompositions of utility dispersion in the economy. First, I decompose utility across firms by firm wage premium, firm amenity and their correlation. I find that amenities drive 75% of the dispersion in utility. This is mainly driven by the fact that my estimates of firm wage premia are quite low. I find that amenities and wage premia are positively correlated, which seems to contradict the CDs model of the labour market when considering the total amenity bundle and not a specific amenity. Second, I include sorting in the picture and decompose between-firms differences adding the contribution of the average worker productivity. "Sorting" is the result of the differential mobility rate of high types that end up higher in the firms-utility-ladder faster. I find significant sorting of workers on both firm premia and unobserved amenities which contribute respectively 8 and 29% to total between-firms utility differences. When I estimate Sorkin's model, I find negative sorting of workers and total compensation which implies that utility is 44% less dispersed. This odd result suggests that not modeling workers heterogeneity might result in very biased estimates whereby productive workers are in bad firms. Third, I provide a novel decomposition to

[3] Déclaration Anuelle des Données Sociales produced by the French national statistical agency

[4] I do not perform bias-correction techniques such as Kline et al. (2020), this is left for future work.

show how CDs attenuate inequality in the labour market by comparing CDs with a counterfactual where firms cannot compensate workers for amenity differences. I find that CDs reduce inequality by 8.1%, as opposed to 100% if they were to offset completely the amenity differences.

My final decomposition looks at how wage premia, amenities and sorting differ across groups. I break down my results across income terciles and age groups. The breakdown by terciles reveals a surprising result. I compute the covariance of wage premia and non-wage values and find that it is negative for the 2 highest income terciles whereas it is robustly positive for the lowest one. This suggests the presence of CDs for workers that are not too low in the income distribution. However, I identify a salient firms-utility-ladder for workers in the lowest tercile. Age groups differences reveal that workers tend to sort towards firms with more amenities as they age, consistent with findings from Lentz et al. (2023).

In order to assess the robustness of my estimates, I use several approaches. First, I bootstrap the entire estimation and document the uncertainty in my estimates. Using methods from Mogstad et al. (2020), I compute the marginal and simultaneous confidence sets for the ranking of firms. Second, I test how the limited mobility might provide too dispersed estimates and find that my results are both qualitatively and quantitatively unchanged across experiments. Lastly, I provide some sensitivity checks to document the direction of potential residual biases.

Literature review - First, I contribute to the literature on compensating differentials. This line of research was fathered by the canonical theory of equalizing differences in Rosen (1983). Since then, economists have struggled to identify CDs in the data because of labour market frictions (Bonhomme and Jolivet 2009 or Lang and Majumdar 2004), see Lavetti (2023) for a recent review. Some recent works that have identified CDs include Lavetti and Schmutte (2018) for fatality risk, Wissmann (2022) and Anelli and Koenig (2022) for health at work, Nagler et al. (2023) for stress and pressure, Chen et al. (2019) and He et al. (2021) for flexible work, Desiere and Walter (2023) for shift work and Dube et al. (2022) for dignity at work. All in all, CDs have been identified in number of settings, but they are very sensitive to frictions (Nagler et al. 2023) and are well identified only at the amenity level. Some other works have identified the absence of CDs: Park et al. (2021) for injury risk and temperatures or Marinescu et al. (2020) for labour rights violations.

Second, my work connects to the broader topic of the total dispersion of utility in the economy and sorting on preferences. Through this lens, wage variations not only compensate differences in amenities but can also augment them or provide rents to workers depending on their preferences. This has been identified in the stated preferences survey of Maestas et al. (2023) where Willingness To Pay (WTP) for amenties vary across groups and is higher for workers that select into jobs with this amenity. With a related approach, Sockin (2021) finds that WTP for job amenities increases with income and augments total compensation inequality. Taber and Vejlin (2020) estimate in a structural model that non-pecuniary preferences interact significantly with other dimension of the labour market such as human capital accumulation and pre-market skills. Lamadon et al. (2022) find in a model without frictions that idiosyncratic preferences play a significant role under-the-hood of wage differences and actually increase total compensation dispersion. Other structural works that identify a significant role for non-pecuniary motives for worker choices of firms include Sullivan and To (2014b), Lehmann (2023) and Lentz et al. (2023). Hall and Mueller (2018) identifies the dispersion of job values from the acceptance choices of unemployed workers. Overall, estimates of the contribution of the non-wage bundle to total compensation inequality are in the range of 5 to 60%. Therefore, there is a lot of uncertainty on the contribution of non-wage amenities to total inequality, but it is robustly positive. Studies that focus on the entirety of the non-wage bundle - as I do - are on the high end of the range compared to the ones that only include a subset. Hence, unobserved amenities seem to actually disperse utility because workers benefit from firm level rents. The utility-ladder theory (as a result of frictions) is often coined as the Mortensen (2005) motive for utility differences.

Third, my work is related to the large literature on sorting in the labour market. In particular, I build on AKM (Abowd et al. 1999) that pioneered fixed effects models with employer-employee datasets. Card et al. (2013) and Song et al. (2019) document how sorting has increased in recent decades exacerbating the differences in wages across firms. Babet et al. (2023) documents a similar pattern in France. Many works since AKM have argued that this estimator is biased and proposed alternative estimation methods, for example Kline et al. (2020) with their leave-one-out correction. Bonhomme et al. (2023) show that across developed economies, unbiased AKM wage variance decompositions lie in the range of 5% for firm effects and $10-20\%$ for sorting of workers.

Finally, I build on the methods from Sorkin (2018) also used in Sorkin (2017). His ranking algorithm is a direct application of the PageRank algorithm developed by Google (1998) to rank web pages by using the citations network and unveil influence on the net. Since then, rankings have much developed and adopted a similar recursive formulation, see Palacios-Huerta and Volij (2004). Bagger and Lentz (2019) also estimate a ranking of firms but in order to identify the productivity job-ladder. Sorkin's method has since received relatively little attention until works by Moser and Morchio (2023) and Lachowska et al. (2023). My work extends this method by introducing heterogeneity and showing the mapping between flow and expected values. However, I resort to approximations and therefore do not make the best use of this technology.

The next section outlines my model, section 3 details the estimation procedure, section 4 contains information on my estimation sample, section 5 presents my results and section 6 concludes. The Appendix contains proofs A, robustness checks B and additional figures and tables C.

2 Model

In this section, I present my framework and show some results about utility dispersion in the economy. The model is a partial equilibrium random on-the-job search model with heterogeneous workers and firms. I first present workers utility and preferences, then move to the reduced-form firms block and finally discuss utility dispersion in the economy and properties of the value functions.

Workers

The economy is populated by N workers that are employed in K firms. Those workers are tenured employees so they are not affected by shocks to their employment status and I also assume that there is no job destruction. Hence, their choice of firm can be seen as a location choice with mobility frictions. A worker (she) derives utility from the wage w her firm pays her and the amenity a of being employed at this firm. The utility derived from the relationship at each period is assumed log-additive:

$$U = \ln w + \ln a$$

Workers discount the future at rate β. At each point in time, workers can relocate. I assume there is some probability λ with which workers can start looking for other opportunities. If they can search, they eventually meet a firm j drawn uniformly from the distribution of firms. I assume uniformity for simplicity and to avoid having an additional need for identifying the offers distribution[5]. Upon meeting, the worker considers the expected utility in both her current firm V_i and the poaching firm V_j and draws a taste shock ϵ for each. She then compares the two alternatives and makes the choice maximizing utility:

$$\max\{\underbrace{V_i + \epsilon}_{\text{stay}}, \underbrace{V_j + \epsilon'}_{\text{move}}\}$$

I make the classical assumption that those shocks are distributed Extreme Value (EV) with variance parameter ν. For now, assume $\nu = 1$. All unobserved drivers of the choice else than wage or the common amenity a are captured by this shock. This can be for example, how close people happen to live to their new workplace. ϵ could be a function of a many-dimensional object summarizing all subjective elements entering the worker's choice of a firm or heterogeneous preference weights. Since I do not observe it, let's assume it is iid across workers and that it captures all the idiosyncratic portion of expected utility.

On the contrary, V_i is the expected value of working at firm i. This is the core object I am interested in. It contains the flow utility given by the individual wage and the amenity which is common to all workers. Arranged in ascending order, the sequence of V_i constitutes a "ranking" of the firms that I can measure in terms of expected utility - I will detail later its expression and link flow and expected utility. Before drawing the idiosyncratic taste shock, all workers have some value for V_i that relies on common fundamentals for which preferences are similar across workers.

The literature is somewhat divided into the proper way to include amenities in labour market models. Lamadon et al. (2022) and Berger et al. (2023) define amenities as the ϵ idiosyncratic preference for firms. This is unobserved to the firm and can therefore be considered as a rent in favor of the worker. On the other side, the literature interested in pricing amenities provided by firms assumes amenities are shared by all workers. This literature is interested into determining whether firm-level differences in amenities explain the firm wage premia (Rosen) or correspond to rents (Mortensen). However, workers could also have heterogeneous values for this bundle. The question of whether there is some systematic component to amenities - net of idiosyncratic differences - is in fine empirical. If people were only to sort on some ϵ, there would be no systematic mobility apart from the one that derives from wage premia differences across firms.

Note that if workers enter the labour market directly in firms that best suit their taste, their mobility would not reveal their preferences. Instead, I assume here that workers initially enter randomly the labour market and then switch firms based on preferences ϵ and total compensation $a + w$. Also, in a model where all amenities are subjective, the welfare implications[6] are quite different than when there is a systematic amenity a. Lowering frictions would mechanically increase welfare because workers can move towards their highest draw ϵ as compared to going towards the firm at the top of the ladder in the shared amenity model.

[5]Sorkin (2018) and Lehmann (2023) identify the offer distribution from the distribution of hires from unemployment. I am not willing to make such an assumption because it would potentially overestimate the recruiting intensity of the worst firms while underestimating the one of the best. Faberman et al. (2022) show that the returns to search and the distribution of offers are very different for employed and unemployed workers.

[6]The welfare statics in discrete choice models is unclear, see Mongey and Waugh (2024)

Firms

There are K firms, which I store in the set Ω. The firms hire workers, produce, pay wages and provide some amenity. I assume firms are characterized by a tuple (ψ, a) of wage premium and amenity. They are both exogenous and I am only interested in estimating their distribution for the economy I observe through the data. Whether they correlate negatively (Rosen) or positively (Mortensen) is ex-ante unknown. I frame my model differently than other papers in the CDs literature, including Sorkin (2018), who instead say firms have heterogeneous costs of amenity-provision. Costs map more easily to the canonical theory of equalizing differences from Rosen (1983). Instead, in my model, amenities are free to the firm and their level is exogenously set for simplicity. Also, wage premia are reduced-form. They are unmodelled but could correspond to productivity differences (Bagger and Lentz 2019), product rents (Wong 2019) or barriers to entry.

Workers are heterogeneous to the firm. I assume there exists a continuum of worker productivity types α in the economy with some distribution F. Production is worker-specific and additive between workers. There are no gains from scale nor complementarity between workers. There is no complementarity between firm and worker types either such that they are both separable. I assume unit-production y is exponential and express profit π as the surplus of production over wage bill:

$$y_i(\alpha) = \exp(\psi_i + \alpha) \qquad 1 + \pi_i(\alpha) = \frac{y_i(\alpha)}{w}$$

In classical models of the labour market, all workers are paid their marginal production - perfect competition - which implies no profits for the firm and $w_i(\alpha) = \exp(\psi_i + \alpha)$. Under this assumption, firms don't discriminate workers based on their productivity because they will in fine not make any profit and rents are completely passed on to wages.

Instead, in order to introduce differential mobility across worker types, I assume that the profit from hiring a worker α is a share μ of its type plus a profit shock ι:

$$1 + \pi_i(\alpha) = \exp(\mu\alpha + \iota) \qquad \mathbb{E}_\iota\Big[\pi_i(\alpha)\Big] = \exp(\mu\alpha) - 1$$

Under this specification, the expected profit is increasing in the worker type α. The profit shock ι can be though of as a match effect. The RHS expression comes from the fact that I assume that ι has mean 0 and that it is distributed logistic[7]. Firms perfectly observe α which is reasonable when considering tenured employees with some employment record. I further assume that none of the match effect is passed on to the worker's wage[8]. Given the proportional profits assumption, the wage of a worker α working in firm i is given by:

$$\ln w_i(\alpha) = (1 - \mu)\alpha + \psi_i$$

The wage premium of the firm appear in the wage - it could be rents of CDs. However, wages only reflect part of the worker productivity advantage because the firm is collecting some of the additional profit of good types (or paying the cost of unproductive workers). With this wage, workers are happier to work at higher ψ firms but in any firm they only get $1 - \mu$ of their productivity. Ex-ante, a firm that meets a worker α will hire her with probability:

$$\rho_\mu(\alpha) \equiv \mathbb{P}(\pi_i(\alpha) > 0) = \mathbb{P}(\mu\alpha + \iota > 0) = \frac{\exp(\mu\alpha)}{1 + \exp(\mu\alpha)}$$

[7] This could be relaxed and calibrated more precisely in the estimation but I stay with this simple version.
[8] This is not necessary either but very convenient

The acceptance probability $\rho_\mu(\alpha)$ for a type α increases in α and increases in second-order stochastic dominance as μ increases. Note that the acceptance probability is independent from i. μ governs the sensitivity of firms to worker types. Hence more (less) productive workers are more (less) often hired as μ increases. The intuition is that, as firms have to bear the cost of bad workers - from my assumption on the relationship between profits and types -, they will only hire them if the profit shock is very good, which occurs less when they bear more of this cost, i.e. higher μ.

μ can be interpreted in many ways. One way to see it may be that there is some labour market power which is uniformly exercised by all firms in the economy. An other microfoundation could be the institutional environment. In what follows, I use the terms firm fixed-effects and wage premia interchangeably.

Utility dispersion

I summarize the model with the following firm i value function for worker α:

$$
\begin{aligned}
V_i(\alpha) = \overbrace{v_i(\alpha)}^{\text{flow value}} + \beta \Bigg[& \overbrace{(1-\lambda) \times V_i(\alpha)}^{\text{no search (1)}} \\
& + \lambda \times \Bigg(\overbrace{(1-\rho(\alpha))V_i(\alpha)}^{\text{not accept. (2)}} \\
& + \rho(\alpha) \underbrace{\sum_{k \in \Omega \setminus i} f_k \, \mathbb{E}_{\epsilon\epsilon'}[\overbrace{\max\{V_i(\alpha)+\epsilon, V_k(\alpha)+\epsilon'\}}^{\text{choice } |\epsilon,\epsilon'}]}_{\text{option value (3)}} \Bigg) \Bigg]
\end{aligned}
$$

Where $V_i(\alpha)$ is the value of being in firm i for a worker α and $v_i(\alpha)$ is the flow utility $U_i(\alpha) = \ln w_i(\alpha) + \ln a_i$. The recursive formulation makes apparent the flow utility and the continuation value discounted by β. Next period, one of 3 events can occur: (1) nothing happens, (2) the worker can search for a job, but the firm doesn't want to recruit her, or (3) she searches and gets an offer from a firm. In cases (1) and (2), she has no choice to make and stays at her firm. However, in case (3) she might move if she meets a good firm V_k and/or draws a good-enough net taste shock $\epsilon' - \epsilon$.

λ governs the frictions in the labour market: how often can workers start looking for an other firm. $\rho(\alpha)$ encodes the acceptance probability for a given type α: how likely are firms to recruit a worker with this productivity level. Upon successful search, workers meet one out of the $K-1$ other firms, sampled uniformly with probability $f_k = f = 1/(K-1)$. The expected value is taken over the 2 taste shocks. Since the acceptance function of the worker and the sampling probability are independent of k, the acceptance probability can be combined into $\lambda\rho(\alpha)$. Note that V is a function of the worker type because different workers are paid differently and have different acceptance probabilities. Revealed preferences appear only in the last term of the value function. From the choices of all the workers in the economy, I will identify the firms ranking and $V_i(\alpha)$s.

Using the properties of the EV distribution, I rewrite the expectation to get:

$$\eta_i(\alpha) \equiv \sum_{k \in \Omega \setminus i} f \log\Big(\exp(V_i(\alpha)) + \exp(V_k(\alpha))\Big)$$

Where $\eta_i(\alpha)$ is the *option value* of the worker. Each firm expected value incorporates the value of the option to move to a better firm. This option is a function of α since Vs depend on α. The option value is always strictly greater than the value of the current firm as can be seen in the formula. However, as workers are higher in the firm-ladder, the option to move to an other firm loses value compared to staying at their current firm. Hence for the highest ranked firm $\eta_i(\alpha) - V_i(\alpha)$ is lowest, but still positive.

I next state 2 propositions that allow me to simplify the value function. *Proposition 1* shows how to remove the worker individual productivity premium and *Proposition 2* how to compare dispersion of utility across types. The proofs Appendix A goes more into the detail of the derivation.

Proposition* 1*: *Under the assumptions that utility takes a log-additive form and that profits are a constant share of α, the ranking of firms according to V^α is the same as with v and $V(\alpha)$ for all α. We can write the value function net of the individual effect as:*

$$V_i^\alpha = v_i + \beta\left[(1-\lambda\rho(\alpha))V_i^\alpha + \lambda\rho(\alpha) f\ \eta_i^\alpha\right]$$

with $v_i = v_i(\alpha) - (1-\mu)\alpha$ *and* $\eta_i^\alpha = \sum_{k\in\Omega\setminus i} \log\Big(\exp(V_i^\alpha) + \exp(V_k^\alpha)\Big)$

Proof: see Appendix A. □

Proposition 1 is the first step towards dealing with the dependence on α in the value function. It defines the alternative value function V^α that removes the worker wage premium from the flow value. Basically, since a worker is paid $(1-\mu)\alpha$ in all firms, this component just shifts all firm expected values up or down with α and can be removed from the value function. Expected values are still functions of α because the option value is weighed by the worker specific transition probability $\rho(\alpha)$. Note that the flow value v is now common to all workers:

$$v_i = v_i(\alpha) - (1-\mu)\alpha = \psi_i + a_i$$

A ranking of firms is a set of value functions that represent an order of preferences between firms. For all αs, *Proposition* 1 says that all workers agree on which firms they find delivers highest and lowest utility. It also states that rankings are similar across all characterisations of utility: the original expected value $V(\alpha)$, the flow value v and the *net-of-person-effect* expected value V^α. This derives from the fact that all workers value the amenity a similarly and all receive the firm rent ψ in their wage. This is a first result that derives from my assumptions. However, these rankings are in different units/scales: small letters v correspond to flow utility whereas capital letters V to expected utility. From here on, I call firm value the common flow value v and V^α the expected value.

The dependence of firm expected values on α appears now only through its effect on mobility. More productive workers will be poached more often and therefore climb the firm-utility-ladder faster. Since they are more likely to be successfully contacted upon searching, they will benefit more from their *option value*. In return, their expected values will be more compressed. On the contrary, low α workers will not be poached very often so their expected values will be more dispersed across firms, even though high and low α workers value ψ and a similarly.

This difference in expected utility dispersion is not trivial. Since I want to measure utility based on choices, I have to carefully measure utility dispersion across firms and across types. This is all the more important because different workers are likely to work in different firms.

Proposition 2*: Under the same set of assumptions as Proposition 1, we have the following identity* $\forall \alpha, i, j$:

$$\Delta V_{ij}^{\alpha} \times h(\alpha, i, j) = \Delta v_{ij}$$

with $h(\alpha, i, j) = 1 - \beta + \beta \, \lambda \, \rho(\alpha) \, (1 - \chi_{ij}^{\alpha})$ *and* χ_{ij}^{α} *the relative difference of option values* η_i^{α} *ranging from 0 when* (i, j) *are the worst firms to 1 when they are the best ones.*

Proof: see Appendix A. □

Proposition 2 characterizes the dispersion of expected utility in the economy as a function of the common flow values. It links flow utility to expected utility. The function h, that takes as arguments the worker type and the location of i and j in the ranking of firms, converts differences in flow utility into α-specific expected value differences. As anticipated, this dispersion is decreasing (h increasing) in α since good workers are more mobile.

The function χ_{ij}^{α} takes as input the position of (i, j) in the ranking of firms and encodes the contribution of the option value to the expected values dispersion. It is quite natural that for firms with the highest expected value, the difference in expected values does not depend on the worker type so much because workers are already at the top of the ladder. Indeed, for high V^{α}s, $(1 - \chi_{ij}) = 0$ so the effect of $\rho(\alpha)$ disappears and $h \approx 1 - \beta$. Basically, the dispersion in expected values at the top of the firm-ladder is equal across worker types. With the notation χ_{ij}, I hint towards the fact that I will approximate[9] χ_{ij}^{α} with some function χ_{ij} in the estimation since they are very close given my estimated F.

Additional remarks

Before moving to identification, some remarks are in order to compare my model with the literature:

- **No dynamics in wages:** My framework is fairly static in the sense that there is no learning (Gibbons et al. 2005), nor bargaining (Postel-Vinay and Robin 2002), nor promotions, nor any dynamic change in the wage apart from mobility. This is to obtain the simple additive wage decomposition that is used in AKM. The model focuses on the amenity component and its interaction with worker heterogeneity and frictions.

- **No unemployment risk:** By removing the unemployment risk, I distance myself from traditional labour market models that investigate the role of the job ladder, with at its bottom unemployment[10]. The job ladder in my model does not feature unemployment risk since all workers already hold a stable tenured job. Again, this setting simplifies the analysis and focuses on between-firms utility dispersion.

- **Proportional profits:** My assumption on firm profits deviates from the homogeneous workers assumption in Sorkin (2018). I provide a simple microfoundation for this deviation in the flavour of the labour market power literature Berger et al. (2022). However any type of model that makes the acceptability of good types higher would fit the

[9]In the proofs Appendix A, I show numerically that this is harmless for this application. It derives from the fact that mobility is quite low for all workers. More discussion in the Appendix A.

[10]For a recent work on job ladder with costly unemployment see Jarosch (2023)

framework. I will later validate this core element of the model in the data. Alternatively, one could also assume that workers have a homogeneous mobility probability but that there is a complementarity between high productivity firms and high types which would induce them to match more often. However, this would lead to sorting in equilibrium which is not the case here.

- **Homogeneous tasks:** I disregard in the model the role of human capital and learning. In fact, I will argue in the empirical section that αs reflect within-occupation productivity differences across workers because I will control for occupational differences. Hence my estimates of individual effects will capture the dimensions of workers efficiency not controlled for in the estimation such as experience, education or individual worker traits.

- **Taste shock variance:** A central parameter not discussed until now is the variance of the taste shock ν. This is a core parameter that will allow to map v and ψ and measure utility across firms in log euros. For now, note that assuming $\nu = 1$ is similar to rescaling my utility as v/ν for any ν. Note that if the taste shock has a high variance, it means that the idiosyncratic part of preferences plays a big role the the mobility of workers as compared to wages and the amenity.

- **Costs of mobility:** My framework assumes no mobility costs. This is fairly common in models of on-the-job search but less so when considering inter-industry mobility. Actually, frictions and mobility costs cannot really be disentangled since they are observationally equivalent to the econometrician. Since the trade Artuç et al. (2010) and development Bryan and Morten (2019) literatures find high costs of switching sectors, my results on the between-sectors differentials might not be so robust, but still they make intuitive sense. More structural models, as Taber and Vejlin (2020) or Dix-Carneiro (2014), take into account the role of human capital in mobility of workers.

- **Compensating differentials:** As in Sorkin (2018), my model features CDs in the form of wage premia paid to offset differences in amenities. But in addition, I will discuss the distribution of amenities not compensated by firm premia, which is more interesting in terms of welfare but more tricky to identify. However, amenities do not enter the firm wage-setting equation and therefore the model is quite reduced-form regarding this aspect. This is because I assume free amenities.

- **Information:** My model assumes that workers perfectly know V_j - utility at the firm they meet - which is a strong assumption since this is a firm workers haven't been at. In reality, workers have necessarily more information about the firm they are at than the firm they go to and these information sets are of different nature. There is no learning or information consideration in my model so I assume this uncertainty is captured by ϵ and plays out symmetrically in the choice.

To conclude this section, I connect my model to Sorkin's. Under his homogeneous workers assumption, all value functions are similar across workers. Also, he does not estimate the flow values because of the computational burden of inverting his value function. I will compare my results to the vs estimated with homogeneous workers but without accounting for shocks and heterogeneous recruiting intensities as he does. He gets a firm value of the following form - keeping the assumption of uniform meeting probabilities:

$$V_j^0 = v_j^0 + \beta \Bigg[(1 - \lambda\rho(0))V_j^0 \\ + \lambda\rho(0) \sum_{k\in\Omega} \log\Big(\exp(V_i^0) + \exp(V_k^0)\Big) \Bigg]$$

Using *Proposition 2*, this yields a simpler h function for the dispersion $h(i,j) = 1 - \beta + \beta\lambda\rho(0)(1-\chi_{ij}^0)$. I denote by $\rho(0)$ the uniform acceptance probability. Hence, h depends only on the rank of (i,j). I now move to my identification strategy.

3 Estimation procedure

The estimation requires access to employer-employee data that contains at least a panel on wages and contract types to filter for tenured employees. From there, it is possible to recover values for α, v and a which fully describe the dispersion of utility and wages in the economy according to my model. As in Sorkin (2018), I use the network of flows of workers across firms to estimate their relative values.

The identification procedure combines estimation and calibration to recover jointly the labour market parameters (λ, μ) and the firm values v. It follows the following steps:

- Recover the individual αs (and ψs) from an AKM-type wage decomposition
- Estimate $\hat{v}$s from all flows of workers across firms by using Sorkin's model or some guess for ρ_μ.
- Compute the parameters $(\hat{\lambda}, \hat{\mu})_n$ implied by the location and mobility of αs across firms and return to the previous step until parameters coincide.

Let me hint towards why I have to do the 2 last estimation steps repeatedly. λ and μ must be calibrated to match the mobility that I observe. Note that ρ is the acceptance rate of workers by firms which I cannot learn directly from raw mobility in the data because I don't observe all the offers. If good workers already are in good firms, they have less chances to find a better firm and to switch firms. Hence the observed mobility for high types is biased downward for estimating their acceptance rate. Therefore, I first need to know the extend of the sorting of workers before estimating ρ, hence the iterative procedure.

Step 1: Wage decomposition

As result of the assumptions presented above, I can decompose the wage into a firm effect and an individual effect. I take my model to be annual and estimate it on data spanning over several years. In order to further interpret my α as an individual "productivity", I enrich the model with an X vector of covariates that captures the job-specific component of wage. Thus, I estimate the following FE model:

$$\ln(w_{it}) = \alpha_i + \psi_{J(i,t)} + X_{it}\beta + \epsilon_{it}$$

where i is a worker, $J(i,t)$ indexes the firm of i at period t, X is a vector of covariates that are worker-specific and ϵ_{it} is random noise conditional on firm and worker types. In the model I presented in the previous section, this last assumption holds. Also, since I assume none of the match effect is passed to the wage, there is no interaction term between the worker and the firm. Covariates in X are characteristics that are likely to covary significantly with wages but not a lot across time, namely a function-layer[11] indicator and a regional dummy.

One strength of my setting is that I restrict myself to the set of firms that are strongly connected by worker flows, thus I have a lot of mobility to identify ψs. This guarantees that my estimates are insulated from the limited mobility bias. On the contrary, since my panel is quite short to ensure stability of ψ and a, my estimates of αs could be less precise. This is somewhat handled by my parsimonious use of those individual workers estimates but instead of percentiles and firm averages. Note that the identification of the AKM model requires at least 2 periods and that all firms are connected by movers. Those conditions are fulfilled here.

Step 2: Identifying values from flows

The second step is to estimate the revealed preferences of workers towards firms from bilateral flows across firms. I build the matrix M of flows for all firms in Ω where M_{ji} is the number of workers going from i to j in my data. Using a first guess that flows of workers do not depend on worker qualities, I use the same model as Sorkin (2018) and express expected flows as:

$$\mathbb{E}[M_{ji}] = \lambda f g_i \rho(0) P(V_j^0 \succ V_i^0) \tag{1}$$

The expected flows are the product of the search probability, the meeting probability, the size of the workforce in the poached firm and the probability that firm j is preferred. I use the notation x^0 to denote the values estimated under the assumption that workers are homogeneous. I rearrange (1) to express $\hat{M}_{ji} = M_{ji}/(\rho(0)g_i)$ and take the ratio of flows between the 2 firms to get the expected relative flows:

$$\frac{\mathbb{E}[\hat{M}_{ji}]}{\mathbb{E}[\hat{M}_{ij}]} = \frac{P(V_j^0 \succ V_i^0)}{P(V_i^0 \succ V_j^0)}$$

This formula intuitively expresses relative flows across firms as a function of the difference in firms expected values. Those probability have a convenient closed form thanks to my assumption on the taste shock distribution: $\exp(V_i^0)/(\exp(V_i^0) + \exp(V_j^0))$. Hence the ratio of probabilities simplifies to:

$$\frac{\mathbb{E}[\hat{M}_{ji}]}{\mathbb{E}[\hat{M}_{ij}]} = \frac{\exp(V_j^0)}{\exp(V_i^0)}$$

Since those flows are taken from the data, there will most likely not be bilateral flows for all firm couples. In order to maximize the information I use in estimating the value of i, I incorporate the inflows from all other firms. I rearrange the previous expression and sum all firms to get:

$$\exp(V_i^0) \underbrace{\sum_j \mathbb{E}[\hat{M}_{ji}]}_{\text{outflows}} = \underbrace{\sum_j \mathbb{E}[\hat{M}_{ij}] \exp(V_j^0)}_{\text{inflows}} \Leftrightarrow \exp(V_i^0) = \sum_j \underbrace{\frac{\mathbb{E}[\hat{M}_{ij}]}{\sum_k \mathbb{E}[\hat{M}_{ki}]}}_{\text{relative flows}} \exp(V_j^0) \tag{2}$$

[11] A function can be though of as broad task that is executed in the firm. Basically production, sales, logistics are functions which can then be decomposed into hierarchical layers.

This equality forms a recursive relationship between all firms expected values using expected flows. Bear in mind that M is a random variable of dimension $K \times K$. To take equation (2) to the data, I use the realization of flows $\hat{M}_{ij}$ and solve for the system $\exp(V^0) = Q\exp(V^0)$ where $\exp(V^0)$ is the stacked vector of values and Q is the matrix as defined by the relative flows coefficients in equation (2). Sorkin (2018) proves that the vector of values can be estimated provided the network of firms is *strongly connected* by flows of workers. Strong connectedness is a property of a directed network which implies that there exists at least one path between all nodes in the network in both directions. In this application, it means I need at least one worker to enter and to leave each firm during the estimation period.

I can solve this system using a power method[12]. To obtain my flow values v^0 from the expected values V^0, I use *Proposition 2* and compute the function h that maps flow values to expected values. I get, for all j, $v_j^0 = \Delta V_{ji}^0 / h(i,j)$ where $v_i^0 = 0$ can be used as a reference firm[13].

Step 3: Estimating differential hiring probabilities

As hinted previously, I need some idea of the firm values to calibrate properly $\lambda\rho$ (jointly). I use the V^0 estimated previously as a starting point. However, since this step will be repeated using V_i^αs I use directly this notation. Let me formally express the model's unconditional probability that a given worker type α moves away from their firm[14]:

$$P(\alpha \text{ moves}) = \sum_k \sum_j \lambda\rho(\alpha) f P(V_j^\alpha \succ V_k^\alpha) H_\alpha(k)$$

Where $H_\alpha(k)$ is the density of workers α that are in firm k. From there, it is straightforward to isolate $\lambda\rho(\alpha)$ and get:

$$\lambda\rho(\alpha) = \frac{P(\alpha \text{ moves})}{f \ \sum_k \sum_j P(V_j^\alpha \succ V_k^\alpha) H_\alpha(k)} \tag{3}$$

The RHS of this equality only contains objects that I can compute from the data and with the previous step results. As my estimation of αs might suffer from imprecision, I group workers in 100 bins before computing the RHS of (3). Since I have 2 parameters in my specification of $\lambda\rho$, I need 2 moments to match. Remember the model formula for $\lambda\rho(\alpha)$:

$$\lambda\rho(\alpha) = \lambda \frac{\exp(\mu\alpha)}{1 + \exp(\mu\alpha)}$$

To match it with some moments $\hat{\omega}$, I run an OLS that gives me the mobility at $\alpha = 0$ with the intercept $\hat{\omega}_0$ and the linear slope $\hat{\omega}_1$. I set $\hat{\lambda} = \hat{\omega}_0 * 2$ (since $\rho(0) = 1/2$) and choose μ to minimize the distance between the slope $\hat{\omega}_1$ and ρ_μ:

$$\hat{\mu} = \arg\min_\mu \mid \sum_\alpha \hat{\lambda}(\rho_\mu(\alpha) - 1/2) - \hat{\omega}_1 \alpha \mid$$

As mentioned previously, the fit can be improved by using an other specification for $\lambda\rho$ but I keep this for simplicity and economic interpretation.

[12]Start with some guess V and update it to QV until convergence.

[13]Indeed this method cannot recover the level of utility but only the dispersion across firms, provided I can measure the taste shock and that utility indeed is log-additive

[14]Note that I cannot say which offers were turned down so I cannot use the destination firms. However, the probability to leave one firm to anywhere else can be estimated by using the distribution of V^αs.

The intuition for this procedure is that taking the raw correlation between mobility and worker types would miss the fact that good workers might already be in good firms. Their raw mobility is therefore biased downward to estimate their acceptance rate. By correcting for their position on the job ladder in (3) I am likely to estimate a steeper acceptance function than the one observed from raw mobility[15]. Ex-ante, I cannot know where workers locate on the job ladder so I have to update my guess for $\lambda\rho$ sequentially.

Step 2 (n): Identifying values from flows of heterogeneous workers

Armed with estimates for individual worker qualities and a guess for the acceptance function, I return to step 2 with heterogeneous workers. This step is repeated with different values for (λ, μ) which will change my estimates for v until convergence. In practice only 2 guesses for (λ, μ) will be enough to obtain a satisfactory match between the acceptance function estimate and the underlying values.

As before, I start by expressing the expected flows from i to j:

$$\mathbb{E}[M_{ji}] = \lambda f g_i \int \rho(\alpha) P(V_j^\alpha \succ V_i^\alpha) dF_i(\alpha)$$

The expected flow of workers from i to j is now the probability that each worker in i: [1] enters the market, [2] meets the firm j, [3] is effectively offered a job and [4] chooses to take it. I take the expectation over the distribution F_i of types working in firm i. Note that H_α was the distribution of type α across firms whereas F_i is the distribution of types within firm i.

As such, this equation is particularly hard to take to the data. Notice that the mobility within the integral is specific to each type. Between α differences in mobility depend both on the differential acceptance rate and on the dispersion in firms expected values. Those 2 forces can push in opposite directions:

- $\partial\rho_\alpha > 0$ the greater the quality of the worker, the highest her chances to get accepted by the poacher.
- $\partial P_\alpha \lesseqgtr 0$ for a given choice between 2 firms, conditional on being accepted, different workers have different probabilities to move. Consider $\Delta V_{ji}^\alpha > 0$. Then $\partial P_\alpha < 0$, the greater (smaller) the quality of the worker, the smaller (greater) will be the difference in firm values ΔV_{ij}^α. This smaller difference pushes "good" workers to move less often than "bad" workers even for slight differences in v - their expected utility is more "compressed". This is because good workers can expect an other match sooner than bad workers. Conversely, with $\Delta V_{ji}^\alpha < 0$, good workers go more often to the lower expected utility firm than bad workers: $\partial P_\alpha > 0$.

I provide a graphical illustration in Appendix Figure 5. It shows that the second force is quite limited compared to the acceptance rate difference.

When considering firm i out-mobility, I integrate over the distribution of its workforce types. If firm i has very unproductive workers, it will actually have less expected outflows than firms that have a higher workforce quality. Therefore, observing a lot of outflows of bad workers is a worst signal of firm quality than seeing a lot of good workers leave it because

[15]The denominator is quite complicated to estimate for all α but much faster when using a partition of the αs space (by binning).

good workers are more attractive to all firms. It is precisely to tackle the workforce composition heterogeneity that I designed this heterogeneity-robust estimator.

If I had enough data, I could estimate all V_j^α within each α or for grouped αs. However, with a mobility rate of around 3% per year within my set of firms, the estimation would suffer from the very limited connectedness of firms through flows of similar α workers.

From this limitation, I propose to evaluate the function $\rho(\alpha)P(V_j^\alpha \succ V_i^\alpha)$ for the average type in the sending firm as a proxy for the integral over the distribution F_i. The next proposition states formally my approximation strategy.

Proposition* 3*: *Within a certain parameters region or when αs are not too dispersed within firms we can use:*

$$\int \rho(\alpha)P(V_j^\alpha \succ V_i^\alpha)dF_i(\alpha) \approx \rho(\alpha_i)P(V_j^{\alpha_i} \succ V_i^{\alpha_i})$$

with $\alpha_i = \int \alpha \, dF_i(\alpha)$

Numerical evidence: see Appendix A. □

Proposition 3 states that I can approximate the average moving rate between 2 firms by its of the average type of the poached firm. In practice, this approximation reduces substantially the nonlinearities present in the model but I gain significantly in tractability and statistical power. Since the integrated function is convex for low regions of the α space and concave at the top, evaluating it at the average point will either decrease the moving probability or increase it respectively. Hence, this approximation potentially exacerbates the effect of types on mobility. I illustrate these patterns more precisely in the Appendix A.

Thanks to *Proposition* 3, I use the flows matrix M as before and express *relative flows* between firms i and j as:

$$\frac{\mathbb{E}[M_{ji}]}{\mathbb{E}[M_{ij}]} = \frac{g_i\ \rho(\alpha_i)}{g_j\ \rho(\alpha_j)} \times \frac{P(V_j^{\alpha_i} \succ V_i^{\alpha_i})}{P(V_i^{\alpha_j} \succ V_j^{\alpha_j})}$$

I rewrite the flows matrix by its quality-weighted equivalent $\tilde{M}_{ji} = M_{ji}/(g_i\ \rho(\alpha_i))$. I further express the probabilities and use *Proposition* 2 to make explicit the dispersion of vs and how it depends on α:

$$\frac{\mathbb{E}[\tilde{M}_{ji}]}{\mathbb{E}[\tilde{M}_{ij}]} = \frac{\exp(\Delta v_{ji}/h(\alpha_i, i, j)))}{1 + \exp(\Delta v_{ji}/h(\alpha_i, i, j))} \times \frac{1 + \exp(\Delta v_{ij}/h(\alpha_j, i, j))}{\exp(\Delta v_{ij}/h(\alpha_j, i, j))}$$

$$\frac{\mathbb{E}[\tilde{M}_{ji}]}{\mathbb{E}[\tilde{M}_{ij}]} = \frac{\exp(v_j/h(\alpha_i, i, j)))}{\exp(v_i/h(\alpha_j, i, j))} \times \frac{\exp(v_i/h(\alpha_j, i, j)) + \exp(v_j/h(\alpha_j, i, j))}{\exp(v_j/h(\alpha_i, i, j)) + \exp(v_i/h(\alpha_i, i, j))} \tag{4}$$

I write $\tilde{h}_{ij} = h(\alpha_i, i, j)$ for notational convenience. What I am after are the values v that best match the data.

Assume for a minute that the last ratio in the product to the RHS - let's call it κ_{ji} - of the previous equation is equal to 1. In this case, the equation reduces to (close to the homogeneous workers case):

$$\frac{\mathbb{E}[\tilde{M}_{ji}]}{\mathbb{E}[\tilde{M}_{ij}]} = \frac{\exp(v_j/\tilde{h}_{ij})}{\exp(v_i/\tilde{h}_{ji})}$$

Then I use the observed flows as a realization of the expected flows to solve for v. I rearrange the previous equation and take the exponent to get:

$$\exp(v_i)\tilde{M}_{ji}^{\tilde{h}_{ji}} = \tilde{M}_{ij}^{\tilde{h}_{ji}}\exp\left(v_j\frac{\tilde{h}_{ji}}{\tilde{h}_{ij}}\right)$$

As in the homogeneous case, I do not observe flows for all pairs of firms. Therefore, to make use of all of the data, I build an estimating equation like equation (2) linking all inflows from and outflows to firm i by summing over all js and then dividing by the outflows:

$$\exp(v_i) = \sum_j \frac{\overbrace{\tilde{M}_{ij}^{\tilde{h}_{ji}}}^{\text{inflows}}}{\underbrace{\sum_k \tilde{M}_{ki}^{\tilde{h}_{ki}}}_{\text{outflows}}} \times \exp\left(v_j\frac{\tilde{h}_{ji}}{\tilde{h}_{ij}}\right) \tag{5}$$

This corresponds to a system of N non-linear equations in $\exp(v_i)$. I stack values in the vector $\exp(v)$ and call h the matrix of $\tilde{h}_{ji}/\tilde{h}_{ij}$s. Using an iteration algorithm, I solve for the vector $\exp(v)$ that satisfies $\exp(v) = \tilde{Q}\exp(vh)$ with $\tilde{Q}$ the matrix of relative flows as in (5).

Now, of course, the ratio κ_{ji} to the right of equation (4) is not 1. Actually, it is the part that governs some of the *dispersion-in-values* effect. Hence, I solve for it by assuming it is 1 and by updating it iteratively until all the objects in the equation have converged. In practice, since $\tilde{h}$ and κ_{ij} are big matrices $(K \times K)$, I try to compute them as little as possible. Still, note that the matrix M is very sparse. In my case, only around 0.4% of the values are non-zero. Hence iterating many times is not so much of an issue compared to computing $\tilde{h}$ which requires to compute the option value in all firms (χ). Still, thanks to the sparsity of my matrices and the closed form of the option value, this is pretty fast. To make the convergence smooth, I use 3 loops to make v, $\tilde{h}$ and κ converge together.[16]

Including κ yields the following equation I am after:

$$\exp(v_i) = \sum_j \frac{\tilde{M}_{ij}^{\tilde{h}_{ji}}}{\sum_k \tilde{M}_{ki}^{\tilde{h}_{ki}}} \times \exp\left(v_j\frac{\tilde{h}_{ji}}{\tilde{h}_{ij}}\right) \times \kappa_{ji}^{\tilde{h}_{ji}} \tag{6}$$

Decomposing the role of the 2 forces I identified previously is not that easy anymore since many terms appear in this equation. The first term corrects for flows of asymmetric worker qualities while the last for the different likelihood of workers to move. The reweighting factor in the exponential function corrects for heterogeneous values dispersion between the workers in the two firms.

This procedure is at the core of the estimation. It links the moments $(\tilde{M})$ to the fundamentals of the economy (v). Note that this model is highly over-parametrized since I have potentially $K \times K$ flows to identify only K firm values. In addition, a small number of moving workers will only provide very noisy estimates. Sorkin bootstraps the estimation by sampling workers without replacement. With his standard errors, he then shrinks his estimates

[16] I do not show that this method delivers a unique solution nor converges. However, my implementation systematically converges to similar solutions for random starting points, which lends confidence in the method.

by correcting for the number of observations for each firm. I do not use this procedure because I keep only a very strongly connected set of firms that guarantees me not too dispersed estimates. However, I still bootstrap the estimation to get standard errors.

In Appendix B, I provide further information on standard errors as computed with the bootstrap and confidence sets for my ranking as proposed in Mogstad et al. (2020). I also perform robustness checks by using different samples. They show that my estimates are noisy but do not seem to overestimate the dispersion in utility. Also, my results are qualitatively and quantitatively similar across tests.

Finally, note that *Proposition 3* is not necessary for estimation. I could instead build the same estimation strategy for individual moves and the associated worker quality by summing over all individual flows. This would capture more precisely the non-linearities in the decision probabilities. However, the estimation of αs is potentially quite noisy (particularly using a 5 years sample as I do) so using the average worker quality at a firm gives more confidence in my estimates. In addition, handling each move individually would make the estimation quite heavy which reduces the computational advantage of my method. Exploring this more granular estimation and ways to improve it is kept for future work.

Convergence of parameters and measuring values

The objective of my estimation is to find v and (λ, μ) that correspond to the vector of worker-qualities $\{\alpha_i\}$, flows of workers M and percentiles of mobility rates. Therefore, I use quite a lot of the data available. This is necessary since I want to recover very granular estimates at the firm level $\{\psi_i, \alpha, v_i\}$. My functional form assumptions grant me an unrestricted estimation for their distribution.

Bear in mind that I implicitly normalized the estimation by using an EV1 specification for the taste shock. Recall that there was absolutely no wage information - apart from the $\rho(\alpha)$ function that maps types to acceptance probability - used in the estimation of vs. As a result, my estimates of flow values v have some unknown scale. Actually, this scale corresponds to taste shock units since I normalized the estimation by using $\nu = 1$. To see this, remember that I set the following worker maximization choice:

$$\max\{V_i^\alpha + \epsilon, V_j^\alpha + \epsilon'\} \quad \Leftrightarrow \quad \max\{\tilde{V}_i^\alpha + \nu\epsilon, \tilde{V}_j^\alpha + \nu\epsilon'\}$$

Now consider a conversion from expected utility to taste shock units given by the parameter ν such that $V_i^\alpha = \tilde{V}_i^\alpha/\nu$ with $\tilde{V}_i^\alpha$ in log euro. In the RHS expression above, I just scaled the expected value by ν which converts both the expected values and the taste shock from taste shock units to the log euro scale (which measures utility by assumption). Hence I also need to identify the scale parameter ν from the data.

Basically, what the estimation process presented above does is to recover $V_i^\alpha = \tilde{V}_i^\alpha/\nu$. I can map it into flow utility using *Proposition 2*: $\Delta V_{ij}^\alpha \times h(\alpha, i, j) = \Delta v_{ij} = \Delta\tilde{v}_{ij}/\nu$. Since the flow values estimation does not use wage information, it is quite natural that it has a scale that depends on how I parametrized the taste shock. This is key to measure well the dispersion of utility.

In order to measure the "scale" of the utility shock in log euro, I compare my estimated Δv with the measurable log euro scale of workers job switches. Hence the link between my taste shock unit estimate and the utility changes in log euro is:

$$\Delta v_{ij} = \Delta\tilde{v}_{ij}/\nu = 1/\nu \times (\Delta\psi_{ij} + \Delta a_{ij})$$

Identifying $1/\nu$ is particularly tricky here because moves result from choices (Roy model) and some unknown underlying distribution for (ψ, a). I present next the logic of my strategy to identify $1/\nu$.

I decompose the amenity into a component that is correlated with ψ and a part orthogonal: $a = \beta\psi + \hat{a}$. If $\beta < 0$, firms paying higher wage premia have worse amenities, consistent with the CDs hypothesis. If I was to make some OLS regression of v on ψ. I would obtain a coefficient that corresponds to $\frac{1+\beta}{\nu}$ due to the fact that I omit a from the regression and that ψ and a are correlated. This basically means that I cannot identify the scale of the taste shock without knowing first the correlation of wage premia and amenities - which is what I am looking for.

Of course this bias would also contaminate a regression of Δv on $\Delta\psi$ because some job switches necessarily depend on unobservable amenities that are somewhat correlated to firm premia. In order to partial-out this effect, I make the assumption that workers that choose firms from unemployment do not factor in amenities in their choice. Hence, they only choose according to $\Delta\psi$ and amenities are orthogonal to their utility choice. The logic is that they might be more financially constrained than Employer-to-Employer (EE) movers and therefore only consider wage differences instead of total compensation. Even if this does not perfectly removes the underlying correlation between amenities and wage premia, it gives me the direction of the bias (whether there are CDs or not). I document my results in the next section.

As a final step, I use the estimates $\hat{v}$ to get the overall dispersion of utility in the economy. These values only matter for their dispersion, hence I normalize them to have a mean 0. I also get the worker-specific values by adding the average worker productivity log euro wage premium: $\hat{v}_i(\alpha_i) = \hat{v}_i + (1-\mu)\alpha_i$. I only stick to firm-level differences and not individual worker utility because my estimates are not necessarily very precise. Of course, I recover the amenity by removing the wage component from the flow values: $a_i = v_i - \psi_i$.

4 Data preparation

I take my model to the French data. I use the Déclaration Anuelle des Données Sociales (DADS) which records all employee-employer relationships each year. The main drawback of this dataset is that it is a repeated cross-section. However, since each file contains detailed information on both year n and $n-1$, it is possible to statistically rebuild the panel version of the dataset. I use Babet et al. (2023) algorithm to do so. This is quite precise in identifying individuals across years[17], therefore I take it as my raw data. This first step is quite important because having a panel is essential to estimate an AKM and therefore the αs.

I restrict my analysis to the period 2010-2015 since I have the best result for the matching algorithm over this period. I estimate my model at an annual frequency. The model is actually quite agnostic about the timing of transitions and could be interpreted as intra-period. However, to estimate the AKM, I need several observations per-worker. I therefore build an annual dataset with one observation per worker per year. In particular, I select each year the main contract of each worker following the Institut National de la Statistique et des Études Économiques (INSEE) methodology[18]. I convert wages in annual wages by using the

[17] I can identify individuals across years for 97.5-99.3% of workers present both years over the period 2010-2015.

[18] Among the contracts that are economically significant, the main contract within the year is the one that payed the highest net wage to the worker. Economically significant employment contracts need to satisfy the 2 following conditions:

	Entire Data	IdF & Restr.	SCS
Observations	63,144,760	10,495,729	4,461,965
Mean wages	9.95	10.30	10.42
Var wages	0.28	0.30	0.32
Skew. wages	0.57	1.02	0.90
Workers	17,593,656	3,176,716	1,383,889
Mean Apparition	3.6	3.3	3.2
Mean empl Dur.	2.7	2.8	3.0
Mean age	39.2	40,5	40.9
Firms	1,189,584	227,556	3,147
Mean Size	15.0	13,7	283.8
10^{th} perc. size	19	17	540
1^{st} perc. size	166	178	2,524
Transitions	9,200,736	825,673	139,304
Mean rate	14.6%	7.87%	3.12%
Share of EE	38.2%	62.2%	68.4%
EE transitions	3,522,209	513,549	95,222

Table 1: Statistics on data sample

number of days worked.

I further reduce the dataset to only keep the Ile-de-France (IdF) région which contains Paris and neighbouring areas. This densely populated region will presumably provide more mobility and more connected firms. I filter for individuals aged between 20 and 63 years old and to keep only employees in private firms (entity code starting by 5). I also remove some sectors that are very unattractive to workers such as interim[19]. Regarding tenure, I only keep individuals that are in a Contrat à Durée Indéterminée (CDI) during year n. I filter for individuals that have an annualized log-wage of at least 8.5 which is already far below the legal minimum. Finally, I only keep full-time contracts because I want the most stable set of employment contracts.

The period I select covers 5 years and includes all transitions during this period. I drop individuals that get out of a full-time contract or to unemployment. As a result, my panel is unbalanced at the worker level because workers disappear if they don't fulfill all the conditions I mentioned. This can be a concern for the AKM estimation but is secondary because I mostly care about having a stable set of firms with enough endogenous job-to-job mobility. In the end, I obtain an average of 3.3 observations per worker which is low but quite expected given the number of filters I apply to the dataset.

In terms of covariates, I have a lot of options with the DADS. My interest is to keep variables that do not evolve too much across time at the individual level but are strong determinants of the wage. This is because my model doesn't speak to promotions or human capital. In particular, I am interested into cleaning the time, occupation and geographic premia from

last for more than 30 days AND 120 hours AND have the ration number of hours / number of days greater than 1.5 OR having paid more than 3 times the minimum wage within the year.

[19] Precisely, I remove NACE A88 codes 01, 02, 03, 78, 87 and 88. Those sectors are sharp outliers in the estimation, with a very low amenity.

the wage. I build a specific dummy for occupation which is the intersection of a function and a hierarchical level. A function is a task that is performed in the firm. It is associated to several occupational categories at all hierarchical levels. This mapping is produced by INSEE[20] so I take it as such. There are a total of 14 functions and 3 layers that I interact together. Examples of functions are logistics, management, transportation or intellectual services.

For the geography, I control for the Départements which is the sub-unit below the région. There are 8 départements in IdF which are very heterogeneous in terms of economic opportunities. Finally, I include year fixed effects which also control for the macroeconomic environment. I purposefully do not include age and sex in the regression because I prefer them to be captured by the αs. This specification is meant to obtain individual fixed-effects that capture most closely within-occupation productive differences. For example, those could be driven by age through experience.

My set of firms is of primary importance. For the second estimation step, I need that firms are *strongly connected* by flows of workers, e.g. that at least one worker enters and leaves each firm. I go further by only keeping firms that have at least 5 workers entering and leaving[21]. Also, I only keep firms that have at least 30 employees at each period and hence appear the entire period. I filter in a loop to insure that all those conditions are satisfied in my final sample and to maintain the strongly connected set.

My final sample of firms is described by Table 1. I provide summary statistics for the raw dataset after applying the panelization alogorithm (Babet et al. 2023) in column 1, after the first filter in column 2 and for the final sample when restricted to the Strongly Connected Set (SCS) of firms in column 3. My final sample contains 1.4 million workers and $3,147$ firms. The wages are much higher than in the full sample because the IdF is the richest region of France. They are also higher than in the sample resulting from the first filter because bigger firms on average pay higher. Workers tend to be slightly older in the final sample compared to the rest of the data. This is certainly due to my selection of tenured contracts.

My final sample of firms is quite small compared to the total number of firms that appear in the DADS. However, they are in the same order of magnitude as INSEE statistics[22] for the firm size distribution even if my analysis is restricted to IdF. The average firm size is 284, the 10^{th} percentile is 540 and the first percentile is $2,524$. However, these firm sizes only record employees working in IdF and therefore underestimate the actual size of firms. Note in passing that the selection of the IdF region will bias the composition of firm workforces compared to the rest of the French economy. This is important to keep in mind when comparing sectors since firms are attached to a sector but not necessarily employ workers in (say) factories in the IdF.

Mobility is central to my model. I define an Employer-to-Nonemployment-to-Employer (ENE) transition as a transition during which the worker has no employment contract for more than 30 days between the main contract of year $n-1$ and the main contract of year n. All other transitions are recorded as Employer-to-Employer (EE) transitions, namely, job changes that do not seem to present involuntary unemployment - defined as more than 30 days without a job. Again, bear in mind that I only use transitions across tenured jobs which are potentially rarer than overall transitions. Indeed, Table 1 shows that going from the raw data to only CDI reduces the transition rate from 15% to 8%. Again, I lose half of my transitions by moving to the more restricted set of firms, most likely because workers entering or

[20]See here

[21]Lehmann (2023) keeps firms that have a total of 5 moves so I am even more conservative

[22]For comparison, according to the INSEE, in 2015, there was 287 firms of more than $5,000$ employees and $5,753$ firms of between 250 and $4,999$ employees.

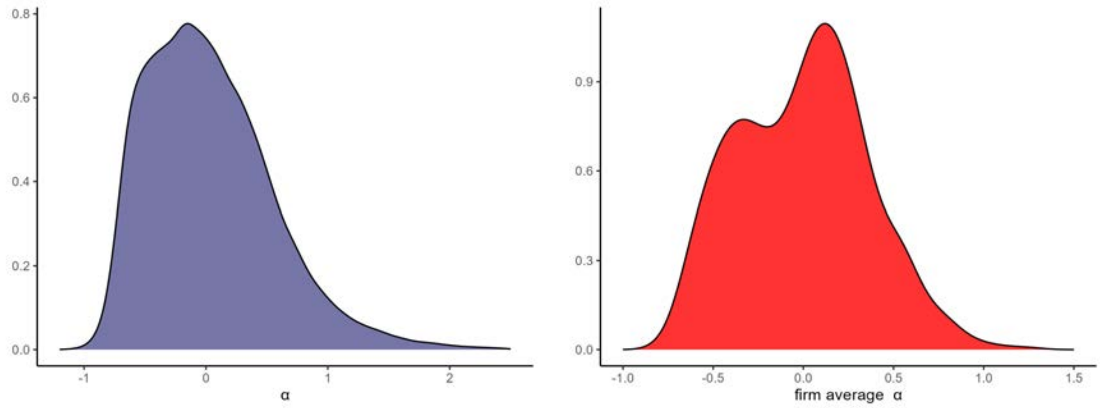

FIGURE 1: Distribution of α and $\bar{\alpha}$

leaving a firm in the SCS are very likely to go in one of the $1,186,437$ French firms not in the SCS considered here. Still, I can count on the $95,222$ EE transitions in the final sample to contain valuable information to discover worker preferences. Note that out of all the transitions I observe in the sample, only 68% are EE transitions that I can use for the estimation.

5 Results

This section presents my results. I first detail the results from the AKM estimation on which the rest of the analysis depends. Second, I discuss the result of the calibration of (ν, λ, μ), the underlying mobility parameters. Third, I dig into the estimated dispersion in flow values and identify the amenity. I leave to Appendix B the analysis of robustness checks and standard errors.

AKM decompostion

The results of the TWFE estimation are given by Table 2. I provide 2 common decompositions for the wage variance. I find that individual effects drive 90% of the dispersion in wages. In comparison, firm effects explain 5% of the variance and this effect is reduced to 2% when removing the covariance with individual effects. Indeed, the correlation of firm and person effects is 0.40 (see Table 12). These results are in line with the rest of the literature. Bonhomme et al. (2023) find that firm effects are in general below 10% when estimated with sufficiently many movers and that the contribution of sorting lies between 10% and 20% in several developed economies. Compared to Babet et al. (2023) who estimate an AKM on the DADS, my results show a lesser contribution of sorting but this might either be due to my sample restriction to IdF or their leave-one-out bias correction (Kline et al. 2020).

In addition, I document in Table 3 the share of variance that is due to within versus between firm differences. I find that half of the variance comes from within the firm, driven by heterogeneous individuals working in the same firm. The other half comes from between firm differences in productivity and sorting. I also report the variance of α for the same within-between decomposition. Indeed, most of the within firms variance comes from individuals, resonating with findings from Song et al. (2019). Additionally, I present in Figure 1 the distribution of the individual effects and of the firms average individual effects since they are a critical input to the estimation. Even if most of the variance in α comes from within firms, a significant portion reflects the fact that different workers work in different firms. Individual effects have a fat right tail and the firm averages distribution seems to be bimodal.

	FE decomp.	Full decomp.
ψ	4.8%	1.9%
α	90%	87.3%
$X\beta$	0.8%	0.4%
2 cov(α,ψ)	-	6.0%
2 cov$(X\beta,\alpha)$	-	-0.7%
2 cov$(X\beta,\psi)$	-	0.0%
R^2	0.96	0.96

Table 2: AKM decomposition

	Within	Between
log w	0.17	0.16
α	0.14	0.14

Table 3: Variance decomposition

Labour market fundamentals

I now present and discuss my estimated mobility parameters and their interpretation through the lens of the model.

In table 4, I report the results from an OLS regression of the individual mobility rate on the α types. All coefficients in Table 4 are significant at all confidence levels. As explained before, I reduce the dimensionality by binning types in 100 bins. The resulting coefficients are given by the first column. I find that the coefficient $\hat{\omega}_1$ of the slope of the observed mobility is 0.008. After recalibrating my observed probability to estimate $\lambda\rho(\alpha)$ instead of $P(\alpha \text{ move})$, I find a higher slope of 0.038. Therefore higher types display less mobility because they already are in good firms. As hinted before, this means that the raw mobility is biased for the acceptance rate of the highest types. I obtain values of $\hat{\lambda} = 0.09$ and $\hat{\mu} = 1.62$.

As can be seen in Figure 2, the fit I obtain does not capture all the non-linearities but performs relatively well. In this figure, the purple dots represent the raw mobility rate as a function of types. The acceptance rate is much steeper when dividing mobility rates by the probability that workers accept offers conditional on their location (in red). The estimated value for μ doesn't have a particular quantitative interpretation in the model. Its scale is the one of the profit shock and is relative to the distribution of αs. It could potentially be rationalised with a more complete model of firm surplus sharing which is not the focus here.

The last parameter is the variance of the taste shock ν that will also allow to convert v in

	P(α move)	$\rho(\alpha)$
$\hat{\omega}_0$	0.021	0.048
$\hat{\omega}_1$	0.008	0.038
λ	-	0.096
μ	-	1.62

Table 4: Mobility regression

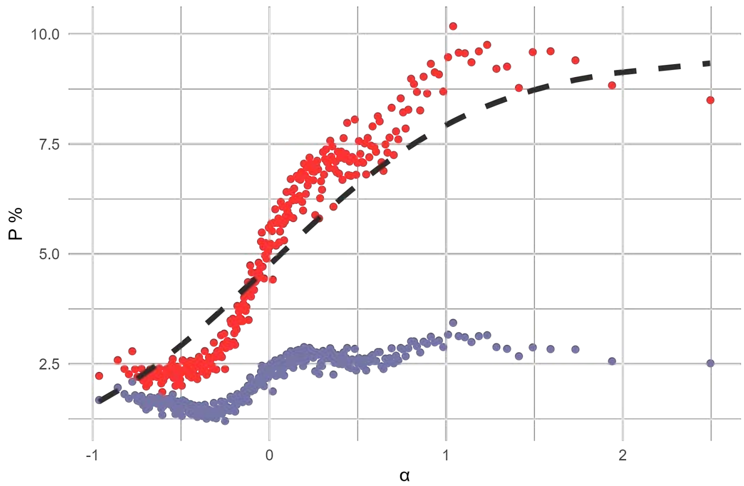

Figure 2: Mobility rates as a function of worker types α and estimated $\rho(\alpha)$

log euro (utility). Recall the formula $v = \tilde{v}/\nu$. To identify it, I compute the Δv_{ij} from realized transitions in the data with the difference in firm premia $\Delta\psi$ associated. I perform an OLS regression of both Δv_0 and Δv on $\Delta\psi$ for the sample of ENE movers and obtain coefficients for $1/\nu$ given in Table 11 in the Appendix. My preferred estimates of $1/\nu$ are 7.1 and 6.6 for the heterogeneous and homogeneous workers models respectively. As explained above, I expect to partial out some of the underlying correlation between ψ and a by using the subset of moves through unemployment. I discuss this approach more in detail in the Appendix B. In addition, I show there how my results are affected by this assumption by varying the underlying unobserved residual correlation between v and a.

Remember that the taste shock enters in the expected value choice such that it cannot be compared directly with the flow utility v. The right formulation in flow utility terms corresponds to the following maximizing choice where $\varepsilon = \epsilon - \epsilon'$ is logistic $(0, 1)$[23] and Δv is in log euro:

$$\max\{ \underbrace{\nu h(\alpha, i, j)\varepsilon}_{\text{net taste shock}} \quad , \quad \underbrace{\Delta v_{ij}}_{\substack{\text{diff. in} \\ \text{flow values}}} \}$$

The h function translates flow utility to expected utility and ν converts the taste shock "units" in log euros. h accounts for the differential dispersion of utility depending on the type of worker and on their location in the job-ladder.

I show in Figure 7 in the Appendix the dispersion of vs and the distribution of the taste shock in log euro for a low-α and a high-α individuals located at the bottom of the ladder. It suggests that the variance of the taste shock is much lower than the overall dispersion in values between firms. Hence, even though idiosyncratic preferences are a strong determinant of workers choices, my estimation finds that they explain little of the choice of workers as compared to the firms-utility-ladder. However, a homogeneous workers model finds that more of the dispersion is explained by idiosyncratic tastes (see LHS of Figure 7).

To illustrate this result, consider the following example. Being offered a 2% annual wage increase keeping amenities constant, a worker of average productivity would accept this offer with probability 78%. This number can go up to 80% when considering a low productivity worker - that cannot expect many other offers in the future - and down to 77% for a high productivity worker. Even though low-productivity workers are more sensitive to utility dif-

[23]The difference of 2 EV1 shocks is distributed logistic (0,1)

	All EE	$\alpha > \bar{\alpha}$	$\alpha < \bar{\alpha}$	All
$P(\Delta w > 0)$	0.57	0.57	0.58	0.55
$P(\Delta \psi > 0)$	0.58	0.59	0.56	0.56
$P(\Delta v_0 > 0)$	0.66	0.66	0.66	0.64
$P(\Delta v > 0)$	0.62	0.64	0.60	0.60
$P(\Delta \bar{\alpha} > 0)$	0.54	-	-	0.54
$P(\Delta size > 0)$	0.58	-	-	0.56
Moves	95,222	56,675	38,547	139,304

TABLE 5: Probability of wage/utility increase upon move

ferences than high productivity workers, this difference in choosing rates is low.

In what follows, I convert v in log euros and discuss the dispersion of utility and amenities across the economy.

Dispersion in utility and amenities

As a first step, I provide evidence that my estimates contain valuable information by comparing them to observables. I then move to the complete analysis of the job ladder and distribution of rents and amenities in the economy.

My first comparison is to validate my model by documenting the likelihood that people move to a higher utility firm upon moving. This is reported in Table 5. First, I find that the probability of having a wage increase upon making an EE move is 57%. This is puzzling that so many job-to-job transitions are to jobs with lower pay. This is the starting point in Sorkin (2018). It potentially hints towards the fact that something unobservable to the econometrician is driving EE transitions, whether it is ϵ or a is an open question. In the second line, I report the probability to move to a higher wage premium firm. It is quite similar if not slightly higher.

When looking at the probability to move to a firm that provides a higher flow value, I find a significantly higher probability which lends credit to the model. For values estimated with homogeneous workers this probability is 66% whereas it is 62% with heterogeneous workers. These are in the range of probabilities estimated by Sorkin. Appendix Figure 11 shows the probability to move to a higher-value firm and the probability to have a wage increase as a function of the firm FE increase $\Delta\psi$ upon moving. It shows that the probability to move to a higher-utility firm is systematically higher even for moves towards firms that pay less. I also report in Table 5 the probability to move to a higher firm and the probability to move to a firm that has a higher average type for comparison.

In addition, I decompose those probabilities for types that are above and below the average type (normalized to 0). In my model, different types have different probabilities to get accepted, but this probability is independent of the firm quality. The only interaction between firm values and worker types comes from the *dispersion in values* effect, namely that low types are more likely to choose to move up because their utility is more dispersed. Therefore the likelihood for a low type to move to a higher utility firm should actually be higher than for high types. This is not what can be seen in Table 5 where high types are 4 percentage points more likely to move to a higher value firm conditional on moving. This could come from several potential sources. One is that there might be complementarities in firms and

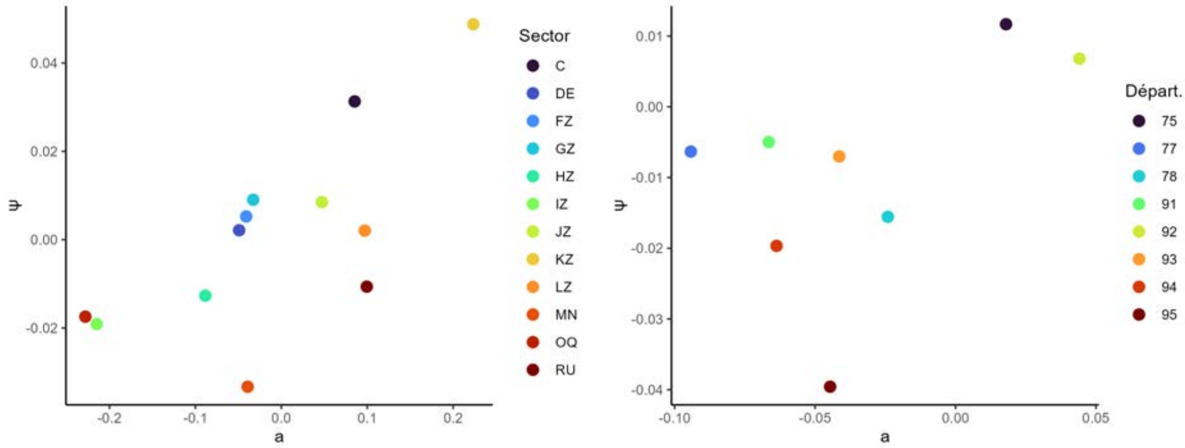

FIGURE 3: Firm average rents and amenities at the sector level (LHS) and by département (RHS)

workers that make the high value firms more likely to hire good workers. An other could be that preferences are non-homothetic[24], namely that lower types have a lower WTP for amenities and therefore prefer moving towards higher paying firms than higher compensation firms. This is hinted by the slightly higher probability of getting a wage increase for low types compared to high types.

Finally, I also include the same statistics for all transitions, including ENE. Intuitively, the probabilities to get a wage or value increase are lower.

Second, I compare my estimates of a and ψ to observables. The first test is to aggregate them at the sector level. I present the employment-weighted values for a subset of 12 sectors classified by the INSEE, for which the correspondence is given in Appendix Table 16. This is shown in Figure 3, left panel. First, note that amenities and wage premia across sectors are positively correlated. This somewhat contradicts the CDs theory of inter-sectoral differences and lends credit to a vertical ladder than spans across sectors. This could also be due to high costs of moving across sectors - inducing more utility dispersion. To prove that my ranking of sectors makes intuitive sense, I quickly compare sectors. In the top right corner, one can find finance (KZ), a high-amenity high-paying sector, and in the bottom left catering and hospitality (IZ) and public administration and health (OQ), tight sectors where employers struggle to hire workers. Notice that manufacturing (C) is quite high in both measures. This could be quite surprising, but keep in mind that I consider IdF activities which are most likely to deliver more amenities than expected in the Manufacturing sector more broadly. Also, bear in mind that that this classification hides very heterogeneous sectors, see detail in Table 14.

The right panel of Figure 3 shows the same aggregation at the Départements level. Paris (75) and Hauts-de-Seine (92) appear with quite high-pay and high-amenity which corresponds indeed to the reality of these economically privileged Départements compared to the others.

I also compute how much of those dimensions explain the total variation in pay premium and amenity in Table 6. Quite surprisingly, even with this coarse classification, a quarter of amenities and wage premia is explained by sectors and Départements respectively. I also use a more granular version of my sectors classification that comprises 68 sectors present in IdF. Reassuringly, I find a strong explanatory power of this grouping for the amenity which

[24]Maestas et al. (2023) and Sockin (2021) find such income-varying preferences for amenities

	ψ	a
Sectors aggregated	10.7%	25.2%
Départements	26.5%	11.5%
Sectors detailed	27.9%	45.0%
Cities	19.1%	24.8%

Table 6: Productivity and amenities with observables

means a captures well the heterogeneity in the non-wage component of sectors. On the contrary, ψs are less explained by this classification - else it is well captured by $X\beta$. This would suggest that there is a lot of within-sector heterogeneity in pay across firms. I provide for illustration in Appendix Table 14 the full ranking of the 57 sectors that contain enough firms to not be identified. It includes both v and v_0 ranks to compare their performance. I also test the more granular city level classification and find that it has more explanatory power for a than Départements whereas it performs worse for ψ. Unfortunately, this classification is less precise because several establishments in a firm can span across cities. Here, I select the city that employs most workers and lose some firms in the process.

Contrary to Sorkin (2018) who uses the LEHD US data, I have access to declared worked hours. This allows me to control whether my individual estimates are driven by differential working times. Since I use annual log wages, my estimates of firm premia contain some worked hours heterogeneity and the amenity potentially captures the satisfaction regarding the number of hours worked. I test for this by regressing my ψ and a on the average weekly hours at the firm level. Average hours correlate positively with both variables which suggests that hours tend to signal a higher-paying firm but also a firm that provides more amenities. I also run the same regression with a quadratic term and find that the positive effect is only concentrated for firms that do most hours. Also, the correlation between ψ and a is unaffected by correcting for weekly hours. See the results for this check in Appendix Table 13.

With confidence in the meaningfulness of my estimates, I provide in Table 7 decompositions for the entire estimated dispersion of utility in the economy. Remember that I can make statements about the dispersion of utility because I converted v in log euro and therefore identify a as the difference between v and ψ. Measurement errors in the estimates of v would lead to a higher role for amenities in the utility dispersion. Again, since I restricted myself to the well-connected set, v is relatively well-identified. In the first decomposition (columns 1-2), I do not include the effect of sorting across firms. This leads to the following decomposition:

$$\text{Var}(v) = \text{Var}(\psi) + \text{Var}(a) + 2 \times \text{Covar}(\psi, a)$$

I find that most of the dispersion comes from unobserved amenities. They contribute 75% to the dispersion of utility as compared to 12% from wage premia. Contrary to the theory of CDs which would imply that the correlation between log euro amenities and log wages is negative, I find a positive correlation between the 2 of 0.21 that contributes a further 12% to overall between firms inequality. For completeness, in Appendix Table 12, I provide the complete correlation matrix for all my estimates. In columns 3-4 of Table 7, I include the effect of sorting on the between firms utility dispersion by using the following decomposition:

$$\text{Var}(v + \bar{\alpha}) = \underbrace{\text{Var}(\psi) + \text{Var}(a) + \times\text{Var}(v)}_{\text{Var}(v)} + \text{Var}(\bar{\alpha}) + \underbrace{2 \times \text{Covar}(a, \bar{\alpha}) + 2 \times \text{Covar}(\psi, \bar{\alpha})}_{2\times\text{Covar}(v,\bar{\alpha})}$$

Note that I use the average worker type in firms $\bar{\alpha}$ to only speak about the between-firms

	no $\bar{\alpha}$	%	incl. $\bar{\alpha}$	%	a_0	%
v	0.065		0.323		0.178	
$\bar{\alpha}$	-		0.136	42%	0.136	76%
ψ	0.008	12%	0.008	2%	0.008	4%
a	0.049	75%	0.049	15%	0.023	13%
$2\times\text{cov}(\psi, a)$	0.008	13%	0.008	3%	-0.006	-3%
$2\times\text{cov}(\bar{\alpha}, \psi)$			0.026	8%	0.026	15%
$2\times\text{cov}(\bar{\alpha}, a)$			0.095	29%	-0.009	-5%

TABLE 7: Dispersion of utility in the economy

differentials[25]. From the AKM results, I hinted towards the fact that workers heterogeneity was the biggest contributor to utility dispersion. Indeed, dispersion in worker types accounts for 42% of utility dispersion across firms. To this, the sorting of good workers to firms with high rents further contributes 8%. All-in-all, observable wage differentials account for 52% of the total utility dispersion. The rest is due to unobservable amenities by 15% and their interaction with the observables. My results suggest a positive utility ladder contributing 3% to utility dispersion. I find that sorting on amenities is even stronger than on firm fixed effects. The correlation between $\bar{\alpha}$ and a is 0.59 and contributes 29% to total utility dispersion.

Columns 5-6 show the same decomposition using a_0 the homogeneous workers estimator. Even though amenities have a high variance and contribute significantly (13%) to overall inequality, they do not reflect sorting nor correlate with firm rents. As a result, the variance of utility in this economy is 44% lower than with v. Actually, the estimated amenities correlate negatively with firm rents -0.24, which pushes inequality down. This is evidence of CDs. Inequality is also reduced thanks to some negative sorting of productive workers on amenities. This result could be driven by the fact that weighting all flows similarly tends to underestimate the actual distance between the top and the bottom of the firm-ladder. Furthermore, notice from Table 12 that the correlation of $\bar{\alpha}$ and v_0 is negative, which means that high-productivity workers are unlikely to work in firms that deliver high utility. This is quite odd but natural given that I use no information on types in the estimation of v_0. Appendix Figure 13 shows a plot of v against v_0 to illustrate how the heterogeneous workers estimator transforms the estimated distribution of values. Hence, neglecting the unobservable heterogeneity in worker productivity leads the homogeneous workers estimator to suggest high CDs but negative sorting of workers on total compensation going against findings from Lamadon et al. (2022) for the US among others.

Coming back to the main estimation, my results imply that including amenities in the welfare analysis increases dispersion in utility across firms by around 8 times. This number is so high only because between firms differences in pay premia are very low once corrected from workers heterogeneity. When including sorting, the number drops to 1.9, i.e. a 47% contribution of unobserved amenities to compensation inequality. Note that this results only applies to tenured workers and not to the entire French labour market. These estimates are somewhat in line with previous findings targeting a similar statistic, higher than Sockin (2021) and Hall and Mueller (2018), in line with Taber and Vejlin (2020), but lower than Lehmann (2023).

To connect with Sorkin's results, I provide the extent to which firm premia compensate

[25] Remember from Table 3 that between-firm differences in pay represent 0.16 of the total variance in wages 0.32. The full analysis of within-firms variance is out of the scope of this work.

for worst amenities. I obtain CDs by regressing ψ on v. The rationale is that CDs are residual pay differences for fixed utility: the part orthogonal to the utility-ladder. On the contrary, the part positively correlated constitutes workers rents at the firm level. As the limit case, consider CDs that perfectly offset utility differences such that corr$(\psi, a) = -1$. In this case, all firms provide the same level of utility: CDs perfectly confound firm premia ψ. The other polar extreme is the case where corr$(\psi, a) = 1$. In this case there is no variation in pay that is orthogonal to v hence no CDs. I find that the R^2 is 53%, which implies that CDs represent 47% of firm premia. Sorkin finds that they represent 70% with his estimates. I find a surprisingly close 66% using v_0. Now consider how this improves welfare in the economy. I can decompose the variation in utility as:

$$\text{Var}(v) = \text{Var}(\text{UL}) - \text{Var}(\tilde{\text{CD}})$$

Where UL stands for Utility-Ladder and is a counterfactual economy where firms cannot compensate workers for differentials in utility and $\tilde{\text{CD}}$ is the part of ψ orthogonal to v. I show this identity in the proofs Appendix A. Note that when corr$(\psi, a) = -1$, Var(UL) = Var($\tilde{\text{CD}}$). This decomposition allows to measure to what extend firms do compensate workers in the economy versus what would be the dispersion in utility if they exercised their full market power (say). I find Var($\tilde{\text{CD}}$) $= 0.0057$ and Var(UL) $= 0.0705$, so firms only compensate 8.1% of the counterfactual total utility dispersion. In other words, without CDs, inequality in the labour market would be 8.1% higher (with v_0 I find 22.4%).

As a final analysis, I provide in Table 8 a decomposition of utility dispersion for different groups in the IdF labour market. Log wages, firm wage premia ψ and firm amenities a are computed across age groups, income terciles, sexes and geography. Wages capture the overall inequality observable trough the raw data. Firm premium and amenity show where groups of workers are located and how this influences their utility. I add the covariance of ψ and a for each group to show how they experience the wage-amenity trade-of.

First, the first 3 rows present results for terciles of the income distribution. Consistent with the previous results, I find that income terciles are ranked positively on the basis of their average firm wage premia and amenity. More paid workers are on average in firms that pay more and provide better non-pay compensation. Interestingly, the variance in amenity decreases with income suggesting that higher-paid workers are more concentrated in high amenity firms. My most striking result is displayed in the last column. I compute the covariance between the firm premium and the amenity across each tercile. Since there is a lot of heterogeneity within firms as seen in Table 3, most income terciles are represented in all firms. Hence, computing the covariance between a and ψ for terciles boils down to computing the same covariance as Table 6 but weighing firms by their share of employment in each tercile. I find that the covariance is negative for most workers except the lower tercile where it is positive. A negative correlation means that firms that pay more provide lower amenities and conversely. It therefore reveals the presence of CDs for most of the income distribution. On the contrary, for the lowest tercile, this relationship is positive, suggesting that differences in amenities go in the same direction as pay differences. This result suggests that most of the utility dispersion due to the positive correlation of wage premia and amenities is driven by low-pay workers for whom the firm-utility-ladder is steep.

I also run the same analysis for worker productivity quantiles and find a similar pattern. Since individual productivity is not observable directly from the data, I prefer to provide my results in terms of observables such as income terciles. In addition, I show in Figure 4 this covariance for more quantiles of the income distribution. Similarly, this covariance is found to decrease with income and turn negative suggesting that CDs are present but not for the

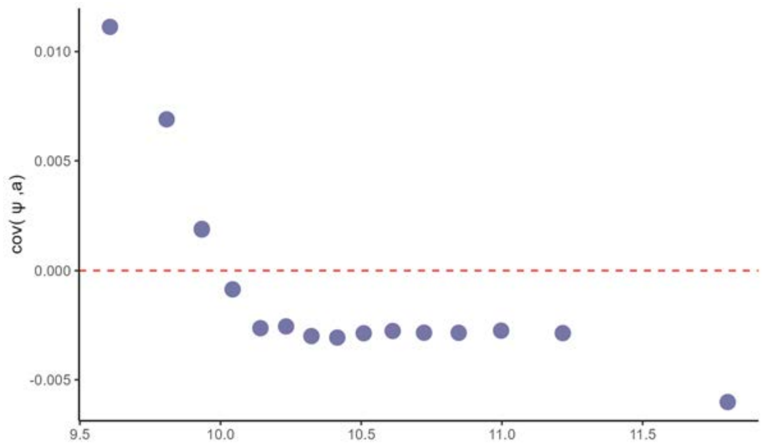

FIGURE 4: Covariance between a and ψ per income quantiles

	Obs.	Trans.	$\bar{w}$	Var(w)	$\bar{\psi}$	Var(ψ)	$\bar{a}$	Var(a)	2cov(ψ, a)
Income terc. 1	301,897	4,450	9.91	0.039	−0.032	0.006	−0.033	0.051	0.007
Income terc. 2	301,875	5,838	10.4	0.019	0.003	0.005	0.109	0.035	−0.003
Income terc. 3	301,864	7,579	11.1	0.185	0.032	0.006	0.156	0.027	−0.003
Age $[22-36]$	303,324	8,719	10.3	0.178	−0.003	0.006	0.056	0.042	0.004
Age $[37-47]$	306,200	5,903	10.5	0.321	0.002	0.006	0.080	0.044	0.004
Age $[48-63]$	296,112	3,245	10.6	0.417	0.003	0.006	0.097	0.046	0.005
Female	306,616	6,036	10.4	0.270	0.008	0.006	0.099	0.043	0.003
Male	599,020	11,831	10.5	0.353	−0.003	0.006	0.066	0.045	0.005
Paris	185,141	5,001	10.6	0.411	0.014	0.008	0.100	0.050	0.010
IdF excl. Paris	720,495	12,866	10.5	0.302	−0.003	0.006	0.071	0.042	0.003

TABLE 8: Dispersion of utility across groups in 2014

lowest rungs of the labour market. Lastly, Figure 7 in the Appendix, which shows the entire distribution of v, shows that the left tail of firm vs is much more dispersed than the right tail, suggesting that the lowest utility firms are further apart in the utility-ladder than higher-up firms. Since those firms are low in the ranking, they certainly benefit from rents from frictions (Mortensen). Since workers are sorted, lower income workers are over-represented in this part of the firm distribution and have less chances to climb up the ladder.

The results across age classes are quite intuitive. As age increases, the average wage increases and sorting accentuates therefore increasing the variance in wages. Similarly, amenities tend to increase with age consistent with other findings that non-pecuniary aspects of jobs gain in importance throughout life (Lentz et al. 2023). These results also suggest a life-cycle dimension whereas workers climb up the utility-ladder during their life, consistent with the mechanisms of my model.

Sexes do not display very significant differences. Women seem to obtain higher amenities but this could be driven by some selection into CDI that is unmodelled here. Indeed, there is quite a big difference in the number of observation between the sexes in my sample, driven by large differences in Full-time and tenured employment for French women[26]. I do not expect my results to be very robust on this front but they give interesting summary statistics. Finally, Paris, as opposed to the rest of the IdF, offers higher firm premia and amenities. In addition,

[26]See here

its job-ladder is much steeper.

6 Discussion and final remarks

In this work, I recover the entire distribution of pay premia ψ, amenities a and worker qualities $\bar{\alpha}$ in the French IdF labour market. Conditional on some structural assumptions, an AKM model allows me to disentangle firm wage premia and worker productivities. I then use Sorkin (2018) structural flows framework, augmented to allow for workers heterogeneity, to tell apart amenities from the total utility delivered by the firms. I estimate parameters pertaining to mobility and the contribution of idiosyncratic preferences to workers' choices.

In line with previous findings (Maestas et al. 2023), my results suggest that unobserved amenities constitute a significant missing variable when studying welfare in the labour market. I find that some of the wage premia that firms pay can be attributed to CDs, but that they represent a very tiny part of the dispersion in utility estimated across firms. Hence, the utility-ladder theory of the labour market seems more appropriate when considering unobserved utility dispersion across firms.

In addition, I document that sorting on amenities is a large contributor to total compensation dispersion, contributing to 30%, almost as much as heterogeneity in worker productivities. Lehmann (2023) also finds a positive sorting on amenities but to a lesser extend. My results resonate with work from Lamadon et al. (2022) who find that high-productivity workers' compensation disproportionately goes through non-wage worker rents. In addition, I find that CDs do appear across all firms but only for firms high enough in the utility-ladder.

Note that my estimates do not actually relate to the entire utility dispersion in the economy. Indeed, I only document between-firms differences which, for wages, correspond to 48% of total dispersion. To account for total utility dispersion through my model, one would have to take a stand on how well workers are sorted on the basis of their idiosyncratic preference ϵ. As in the broader analysis of CDs (Lavetti 2023), this would run into the problem of disentangling individual preferences and amenity prices, both non-linear objects that interact together. This is outside of the scope of this paper where I only identify the common amenity at the firm level. Taking this firm-level approach allows me to speak to the nexus utility-ladder versus CDs, keeping aside considerations on preferences sorting and equilibrium amenity provision. This also implies that my results does not completely map with work from Taber and Vejlin (2020) and Hall and Mueller (2018).

Note that frictions do not bias my estimation in this model, they actually permit identification. The only reason why workers do not work in their favorite firm is because they cannot relocate at any time. Hence mobility, thanks to frictions, allows to reveal preferences. Keep in mind that I defined sorting as a state of the economy and not an equilibrium outcome. Since bad workers cannot relocate as easily as good ones, they are stuck in bad firms for longer. However, with enough time, they eventually climb the ladder as do productive workers. This is in contrast with models of assortative matching where sorting arises in equilibrium (Shimer and Smith 2000). Here, I am agnostic about such complementarity because I don't build an equilibrium. The underlying assumption is that frictions are a stronger force than complementarity which allows me to have a function ρ independent of the firm value. A more complex model might include such complementarity, which would potentially increase inequality because bad workers are altogether not a target of high value firms.

Sorkin (2018) argues that it is hard to see how worker heterogeneity would bias his estimates of firm values. Indeed, the model is designed such that ϵ captures all unmodeled determinants of choices and mobility. Only systematic biases in mobility or preferences could alter estimates of V. I show that including differential poaching rates across types allows to qualitatively reverse the results. My model generates a more dispersed lower end of the utility ladder where bad workers are stuck for longer. When I further decompose CDs across the income distribution, I find that CDs represent 45% of the wage differentials at the bottom but 70% past the 3^{rd} decile of income as in Sorkin (2018). By reweighing flows of workers by their quality, my estimator delivers similar results as Sorkin's but uncovers a low rung to the labour market where amenity and wages are jointly lower. Not accounting for heterogeneity leads to estimates that imply a negative sorting of workers on the firms-utility-ladder, a quite unattractive result.

An other bias could stem from systematically different preferences for amenities - as a function of location. In particular, non-homothetic preferences imply that workers with low income derive less value for amenities than workers with high incomes. Their choices would in turn reflect more closely pay differences than utility differences. This would push the total values of firms down compared to their true value. Hence non-homothetic preferences in the direction suggested by Maestas et al. (2023) and Sockin (2021)[27] would imply my estimates are too dispersed. Still, only using tenured workers likely reduces the effect of non-homotheticity because they are more secured to look at the total compensation bundle while considering jobs. Again, this caveat pushes towards interpreting my results as an upper bound on utility dispersion.

Note that my estimates for mobility parameters depend on the assumption that mobility is costless. Including costs, in the form of human capital acquisition or dis-utility would imply that the observed mobility is too low compared to potential mobility. This would not necessarily affect the ranking of firms but could affect the measure of the taste shock. Unobserved mobility costs would imply that my estimates are too compressed compared to the actual inequality in the labour market. To wrap up, accounting for these types of forces that systematically push workers to move more or less is an interesting avenue for future research that aims at refining the Sorkin estimator for revealed preferences.

Compared to Sorkin (2018), I keep the matching process quite simple. In his full model, he allows firms to have heterogeneous recruiting intensities. This way, big firms can devote more resources to hiring and have more chances to meet a worker with a high idiosyncratic draw. I do not correct for the job offers distribution being potentially more skewed towards some firms. Again, this derives from my use of tenured workers. The intuition is that on-the-job tenured searchers might not meet firms following the common job postings process as postulated in Sorkin (2018). I assume that workers know the entire distribution of firms and can only start looking for a job from time to time. Indeed, tenured workers are more likely to direct their search towards better firms (Banfi and Villena-Roldán 2019), a topic with still limited evidence (Faberman et al. 2022). Lehmann (2023) finds that his results are not greatly affected by modifying the offers process.

My results are estimated using only tenured employees. Under the assumption that the amenity is common to all employees within a firm, not only tenured ones, my results can be extended to them. Restricting myself to tenured workers is only a some strategy to avoid potential biases stemming from involuntary mobility. Actually, Table 10 even suggests that using also CDD workers delivers a lower contribution of amenities to between-firms utility dispersion.

My assumption on profits allows me to obtain a simple reduced-form acceptance probability. In addition, with additively separable utility and free firm-level amenity, I can fully

[27] They find a higher WTP for amenities as workers have a higher income.

separate the amenity-provision from the wage-setting process. This gives me the AKM framework with a firm wage premium that I can take directly to the total firm flow value. In this simpler setting, Sorkin (2018) shows that this allows to identify CDs and I further show how to properly measure the amenity. Building a more intricate framework is far beyond the scope of this work but could allow to exploit the interesting interaction between workers heterogeneity, market power and amenities in general equilibrium as Berger et al. (2023) and Lamadon et al. (2022) pave the way for. As a result one could exploit how surplus is shared in the presence of idiosyncratic and common amenities.

Lastly, my method to identify the scale of the taste shock is not flawless. Further work could potentially improve it by using an instrumental variable approach or identifying jointly (ν, β) using MLE. In fine, getting a clear idea of the extend to which ϵ contributes to utility is a very important empirical question. Given that Taber and Vejlin (2020) find that total wages would increase by about 18 percent if CDs were to be removed, the extend to which unobserved worker rents pertain to common amenities (inequality) or individual preferences (sorting) is capital to make welfare statements.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (Mar. 1999). "High Wage Workers and High Wage Firms". en. In: *Econometrica* 67.2, pp. 251–333. ISSN: 0012-9682, 1468-0262. DOI: 10.1111/1468-0262.00020. URL: http://doi.wiley.com/10.1111/1468-0262.00020 (visited on 04/28/2024).

Anelli, Massimo and Felix Koenig (2022). "Willingness to Pay for Workplace Safety". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.4114423. URL: https://www.ssrn.com/abstract=4114423 (visited on 05/01/2024).

Artuç, Erhan, Shubham Chaudhuri, and John McLaren (June 2010). "Trade Shocks and Labor Adjustment: A Structural Empirical Approach". en. In: *American Economic Review* 100.3, pp. 1008–1045. ISSN: 0002-8282. DOI: 10.1257/aer.100.3.1008. URL: https://pubs.aeaweb.org/doi/10.1257/aer.100.3.1008 (visited on 04/28/2024).

Babet, Damien, Olivier Godechot, and Marco G Palladino (2023). "In the Land of AKM: Explaining the Dynamics of Wage Inequality in France". en. In: *INSEE Working Paper*.

Bagger, Jesper and Rasmus Lentz (Jan. 2019). "An Empirical Model of Wage Dispersion with Sorting". en. In: *The Review of Economic Studies* 86.1, pp. 153–190. ISSN: 0034-6527, 1467-937X. DOI: 10.1093/restud/rdy022. URL: https://academic.oup.com/restud/article/86/1/153/4995193 (visited on 04/28/2024).

Banfi, Stefano and Benjamín Villena-Roldán (July 2019). "Do High-Wage Jobs Attract More Applicants? Directed Search Evidence from the Online Labor Market". en. In: *Journal of Labor Economics* 37.3, pp. 715–746. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/702627. URL: https://www.journals.uchicago.edu/doi/10.1086/702627 (visited on 05/15/2024).

Berger, David, Kyle Herkenhoff, Andreas Kostøl, and Simon Mongey (Apr. 2023). *An Anatomy of Monopsony: Search Frictions, Amenities and Bargaining in Concentrated Markets*. en. Tech. rep. w31149. Cambridge, MA: National Bureau of Economic Research, w31149. DOI: 10.3386/w31149. URL: http://www.nber.org/papers/w31149.pdf (visited on 04/28/2024).

Berger, David, Kyle Herkenhoff, and Simon Mongey (Apr. 2022). "Labor Market Power". en. In: *American Economic Review* 112.4, pp. 1147–1193. ISSN: 0002-8282. DOI: 10.1257/aer.20191521. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20191521 (visited on 05/04/2024).

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler (Apr. 2023). "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" en. In: *Journal of Labor Economics* 41.2, pp. 291–322. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/720009. URL: https://www.journals.uchicago.edu/doi/10.1086/720009 (visited on 04/28/2024).

Bonhomme, Stéphane and Grégory Jolivet (Aug. 2009). "The pervasive absence of compensating differentials". en. In: *Journal of Applied Econometrics* 24.5, pp. 763–795. ISSN: 0883-7252, 1099-1255. DOI: 10.1002/jae.1074. URL: https://onlinelibrary.wiley.com/doi/10.1002/jae.1074 (visited on 04/28/2024).

Bryan, Gharad and Melanie Morten (Oct. 2019). "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia". en. In: *Journal of Political Economy* 127.5, pp. 2229–2268. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/701810. URL: https://www.journals.uchicago.edu/doi/10.1086/701810 (visited on 05/04/2024).

Card, David, Jörg Heining, and Patrick Kline (Aug. 2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality*". en. In: *The Quarterly Journal of Economics* 128.3, pp. 967–1015. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjt006. URL: https://academic.oup.com/qje/article/128/3/967/1848785 (visited on 04/28/2024).

Chen, M. Keith, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen (Dec. 2019). "The Value of Flexible Work: Evidence from Uber Drivers". en. In: *Journal of Political Economy*

127.6, pp. 2735–2794. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/702171. URL: https://www.journals.uchicago.edu/doi/10.1086/702171 (visited on 05/01/2024).

Desiere, Sam and Christian Walter (2023). "The Shift Premium: Evidence from a Discrete Choice Experiment". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.4580837. URL: https://www.ssrn.com/abstract=4580837 (visited on 05/01/2024).

Dix-Carneiro, Rafael (2014). "Trade Liberalization and Labor Market Dynamics". en. In: *Econometrica* 82.3, pp. 825–885. ISSN: 0012-9682. DOI: 10.3982/ECTA10457. URL: http://doi.wiley.com/10.3982/ECTA10457 (visited on 05/14/2024).

Dube, Arindrajit, Suresh Naidu, and Adam Reich (Sept. 2022). *Power and Dignity in the Low-Wage Labor Market: Theory and Evidence from Wal-Mart Workers*. en. Tech. rep. w30441. Cambridge, MA: National Bureau of Economic Research, w30441. DOI: 10.3386/w30441. URL: http://www.nber.org/papers/w30441.pdf (visited on 05/01/2024).

Faberman, R. Jason, Andreas I. Mueller, Aysegül Sahin, and Giorgio Topa (2022). "Job Search Behavior Among the Employed and Non-Employed". en. In: *Econometrica* 90.4, pp. 1743–1779. ISSN: 0012-9682. DOI: 10.3982/ECTA18582. URL: https://www.econometricsociety.org/doi/10.3982/ECTA18582 (visited on 04/28/2024).

Gibbons, Robert, Lawrence F. Katz, Thomas Lemieux, and Daniel Parent (Oct. 2005). "Comparative Advantage, Learning, and Sectoral Wage Determination". en. In: *Journal of Labor Economics* 23.4, pp. 681–724. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/491606. URL: https://www.journals.uchicago.edu/doi/10.1086/491606 (visited on 05/04/2024).

Google (1998). "The PageRank Citation Ranking : Bringing Order to the Web". In.

Hall, Robert E. and Andreas I. Mueller (Aug. 2018). "Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values". en. In: *Journal of Political Economy* 126.4, pp. 1594–1637. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/697739. URL: https://www.journals.uchicago.edu/doi/10.1086/697739 (visited on 04/28/2024).

He, Haoran, David Neumark, and Qian Weng (July 2021). "Do Workers Value Flexible Jobs? A Field Experiment". en. In: *Journal of Labor Economics* 39.3, pp. 709–738. ISSN: 0734-306X, 1537-5307. DOI: 10.1086/711226. URL: https://www.journals.uchicago.edu/doi/10.1086/711226 (visited on 05/01/2024).

Jarosch, Gregor (2023). "Searching for Job Security and the Consequences of Job Loss". en. In: *Econometrica* 91.3, pp. 903–942. ISSN: 0012-9682. DOI: 10.3982/ECTA14008. URL: https://www.econometricsociety.org/doi/10.3982/ECTA14008 (visited on 05/01/2024).

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2020). "Leave-Out Estimation of Variance Components". en. In: *Econometrica* 88.5, pp. 1859–1898. ISSN: 0012-9682. DOI: 10.3982/ECTA16410. URL: https://www.econometricsociety.org/doi/10.3982/ECTA16410 (visited on 04/30/2024).

Krueger, Alan B. and Lawrence H. Summers (Mar. 1988). "Efficiency Wages and the Inter-Industry Wage Structure". In: *Econometrica* 56.2, p. 259. ISSN: 00129682. DOI: 10.2307/1911072. URL: https://www.jstor.org/stable/1911072?origin=crossref (visited on 05/03/2024).

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen Woodbury (May 2023). *Work Hours Mismatch*. en. Tech. rep. w31205. Cambridge, MA: National Bureau of Economic Research, w31205. DOI: 10.3386/w31205. URL: http://www.nber.org/papers/w31205.pdf (visited on 05/01/2024).

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (Jan. 2022). "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market". en. In: *American Economic Review* 112.1, pp. 169–212. ISSN: 0002-8282. DOI: 10.1257/aer.20190790. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20190790 (visited on 04/28/2024).

Lang, Kevin and Sumon Majumdar (Nov. 2004). "THE PRICING OF JOB CHARACTERISTICS WHEN MARKETS DO NOT CLEAR: THEORY AND POLICY IMPLICATIONS*".

en. In: *International Economic Review* 45.4, pp. 1111–1128. ISSN: 0020-6598, 1468-2354. DOI: 10.1111/j.0020-6598.2004.00299.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.0020-6598.2004.00299.x (visited on 05/04/2024).

Lavetti, Kurt (Aug. 2023). "Compensating Wage Differentials in Labor Markets: Empirical Challenges and Applications". en. In: *Journal of Economic Perspectives* 37.3, pp. 189–212. ISSN: 0895-3309. DOI: 10.1257/jep.37.3.189. URL: https://pubs.aeaweb.org/doi/10.1257/jep.37.3.189 (visited on 05/01/2024).

Lavetti, Kurt and Ian M Schmutte (2018). *Estimating Compensating Wage Differentials with Endogenous Job Mobility*. en. Tech. rep. Draft. URL: https://www.bc.edu/content/dam/bc1/schools/mcas/economics/pdf/seminars/CDEM_v21.pdf (visited on 04/28/2024).

Lehmann, Tobias (2023). "Non-Wage Job Values and Implications for Inequality". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.4668593. URL: https://www.ssrn.com/abstract=4668593 (visited on 04/28/2024).

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin (2023). "The Anatomy of Sorting—Evidence From Danish Data". en. In: *Econometrica* 91.6, pp. 2409–2455. ISSN: 0012-9682. DOI: 10.3982/ECTA16425. URL: https://www.econometricsociety.org/doi/10.3982/ECTA16425 (visited on 04/28/2024).

Maestas, Nicole, Kathleen J. Mullen, David Powell, Till Von Wachter, and Jeffrey B. Wenger (July 2023). "The Value of Working Conditions in the United States and Implications for the Structure of Wages". en. In: *American Economic Review* 113.7, pp. 2007–2047. ISSN: 0002-8282. DOI: 10.1257/aer.20190846. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20190846 (visited on 05/01/2024).

Marinescu, Ioana Elena, Yue Qiu, and Aaron J. Sojourner (2020). "Wage Inequality and Labor Rights Violations". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3673495. URL: https://www.ssrn.com/abstract=3673495 (visited on 05/15/2024).

Mogstad, Magne, Joseph Romano, Azeem Shaikh, and Daniel Wilhelm (Mar. 2020). *Inference for Ranks with Applications to Mobility across Neighborhoods and Academic Achievement across Countries*. en. Tech. rep. w26883. Cambridge, MA: National Bureau of Economic Research, w26883. DOI: 10.3386/w26883. URL: http://www.nber.org/papers/w26883.pdf (visited on 05/13/2024).

Mongey, Simon and Michael Waugh (Feb. 2024). *Discrete Choice, Complete Markets, and Equilibrium*. en. Tech. rep. w32135. Cambridge, MA: National Bureau of Economic Research, w32135. DOI: 10.3386/w32135. URL: http://www.nber.org/papers/w32135.pdf (visited on 05/01/2024).

Mortensen, Dale (2005). *Wage dispersion: why are similar workers paid differently?* eng. 1. MIT Press paperback ed. Zeuthen lecture book series. Cambridge, Mass.: MIT Press. ISBN: 978-0-262-63319-2 978-0-262-13433-0.

Moser, Christian and Iacopo Morchio (2023). "The Gender Gap: Micro Sources and Macro Consequences". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3176868. URL: https://www.ssrn.com/abstract=3176868 (visited on 05/01/2024).

Nagler, Markus, Johannes Rincke, and Erwin Winkler (July 2023). "High-Pressure, High-Paying Jobs?" en. In: *Review of Economics and Statistics*, pp. 1–45. ISSN: 0034-6535, 1530-9142. DOI: 10.1162/rest_a_01362. URL: https://direct.mit.edu/rest/article/doi/10.1162/rest_a_01362/116980/High-Pressure-High-Paying-Jobs (visited on 05/01/2024).

Palacios-Huerta, Ignacio and Oscar Volij (May 2004). "The Measurement of Intellectual Influence". en. In: *Econometrica* 72.3, pp. 963–977. ISSN: 0012-9682, 1468-0262. DOI: 10.1111/j.1468-0262.2004.00519.x. URL: http://doi.wiley.com/10.1111/j.1468-0262.2004.00519.x (visited on 04/28/2024).

Park, Jisung, Nora M. C. Pankratz, and A. Behrer (2021). "Temperature, Workplace Safety, and Labor Market Inequality". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3892588. URL: https://www.ssrn.com/abstract=3892588 (visited on 05/01/2024).

Postel-Vinay, Fabien and Jean-Marc Robin (Nov. 2002). "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity". en. In: *Econometrica* 70.6, pp. 2295–2350. ISSN: 0012-9682, 1468-0262. DOI: 10.1111/j.1468-0262.2002.00441.x. URL: http://doi.wiley.com/10.1111/j.1468-0262.2002.00441.x (visited on 05/04/2024).

Rosen, Sherwin (July 1983). *The Equilibrium Approach to Labor Markets*. en. Tech. rep. w1165. Cambridge, MA: National Bureau of Economic Research, w1165. DOI: 10.3386/w1165. URL: http://www.nber.org/papers/w1165.pdf (visited on 04/28/2024).

Shimer, Robert and Lones Smith (Mar. 2000). "Assortative Matching and Search". en. In: *Econometrica* 68.2, pp. 343–369. ISSN: 0012-9682, 1468-0262. DOI: 10.1111/1468-0262.00112. URL: http://doi.wiley.com/10.1111/1468-0262.00112 (visited on 05/16/2024).

Sockin, Jason (2021). "Show Me the Amenity: Are Higher-Paying Firms Better All Around?" en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3957002. URL: https://www.ssrn.com/abstract=3957002 (visited on 05/01/2024).

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter (Feb. 2019). "Firming Up Inequality*". en. In: *The Quarterly Journal of Economics* 134.1, pp. 1–50. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjy025. URL: https://academic.oup.com/qje/article/134/1/1/5144785 (visited on 04/28/2024).

Sorkin, Isaac (May 2017). "The Role of Firms in Gender Earnings Inequality: Evidence from the United States". en. In: *American Economic Review* 107.5, pp. 384–387. ISSN: 0002-8282. DOI: 10.1257/aer.p20171015. URL: https://pubs.aeaweb.org/doi/10.1257/aer.p20171015 (visited on 05/01/2024).

Sorkin, Isaac (Aug. 2018). "Ranking Firms Using Revealed Preference*". en. In: *The Quarterly Journal of Economics* 133.3, pp. 1331–1393. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjy001. URL: https://academic.oup.com/qje/article/133/3/1331/4813638 (visited on 04/28/2024).

Sullivan, Paul and Ted To (2014a). "Job Dispersion and Compensating Wage Differentials". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2506693. URL: http://www.ssrn.com/abstract=2506693 (visited on 05/01/2024).

Sullivan, Paul and Ted To (2014b). "Search and Nonwage Job Characteristics". en. In: *Journal of Human Resources* 49.2, pp. 472–507. ISSN: 1548-8004. DOI: 10.1353/jhr.2014.0013. URL: http://muse.jhu.edu/content/crossref/journals/journal_of_human_resources/v049/49.2.sullivan.html (visited on 04/28/2024).

Taber, Christopher and Rune Vejlin (2020). "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market". en. In: *Econometrica* 88.3, pp. 1031–1069. ISSN: 0012-9682. DOI: 10.3982/ECTA14441. URL: https://www.econometricsociety.org/doi/10.3982/ECTA14441 (visited on 04/28/2024).

Wissmann, Daniel (Jan. 2022). "Finally a Smoking Gun? Compensating Differentials and the Introduction of Smoking Bans". en. In: *American Economic Journal: Applied Economics* 14.1, pp. 75–106. ISSN: 1945-7782, 1945-7790. DOI: 10.1257/app.20180077. URL: https://pubs.aeaweb.org/doi/10.1257/app.20180077 (visited on 05/01/2024).

Wong, Horng Chern (2019). "Wage Determination Across Firms". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3446088. URL: https://www.ssrn.com/abstract=3446088 (visited on 04/28/2024).

A Proofs

For convenience, I repeat some definitions here:

$$V_i(\alpha) = v_i(\alpha) + \beta\Bigg[(1-\lambda)\times V_i(\alpha) + \lambda\times\Bigg((1-\rho(\alpha))V_i(\alpha) + \rho(\alpha)\sum_{k\in\Omega\setminus i} f\,\mathbb{E}_{\epsilon\epsilon'}[\,\max\{V_i(\alpha)+\epsilon, V_k(\alpha)+\epsilon'\}\,]\Bigg)\Bigg] \tag{VF}$$

$$v_i = \psi_i + a_i \qquad\qquad v_i(\alpha) = (1-\mu)\alpha + \psi_i + a_i$$

$$\eta_i^\alpha = \sum_{k\in\Omega\setminus i} f\,\mathbb{E}_{\epsilon\epsilon'}[\,\max\{V_i^\alpha+\epsilon, V_k^\alpha+\epsilon'\}\,] \tag{OV}$$

$$V_i^\alpha = v_i + \beta\Big[(1-\lambda\rho(\alpha))\times V_i^\alpha + \lambda\rho(\alpha) f\,\eta_i^\alpha\,\Big] \tag{MVF}$$

The goal of this Appendix is to unpack the value function and make proofs as I go. I show *Proposition 1*, *Proposition 2* and then how I approximate χ and *Proposition 3*. A last paragraph shows the variance decomposition with UL and $\tilde{CD}$.

Proposition* 1*: Under the assumptions that utility takes a log-additive form and that profits are a constant share of α, the ranking of firms according to V^α is the same as with v and $V(\alpha)$ for all α. We can write the value function net of the individual effect as:

$$V_i^\alpha = v_i + \beta\Bigg[(1-\lambda\rho(\alpha))V_i^\alpha + \lambda\rho(\alpha) f\,\eta_i^\alpha\,\Bigg]$$

with $v_i = v_i(\alpha) - (1-\mu)\alpha$ *and* $\eta_i^\alpha = \sum_{k\in\Omega\setminus i} \log\Big(\exp(V_i^\alpha) + \exp(V_k^\alpha)\Big)$

From our assumption on the firm's profits, the expression of $v_i(\alpha)$ is as above and we can express the common v_i as in *Proposition 1*. Starting from the original value function (VF), we guess that $V_i(\alpha) = \frac{1-\mu}{1-\beta}\alpha + V_i^\alpha$ and verify:

$$\frac{1-\mu}{1-\beta}\alpha + V_i^\alpha = v_i + (1-\mu)\alpha + \beta\Bigg[(1-\lambda\rho(\alpha))\times\Big(\frac{1-\mu}{1-\beta}\alpha + V_i^\alpha\Big) \lambda\rho(\alpha)\sum_k f\,\mathbb{E}_{\epsilon\epsilon'}[\,\max\{\frac{1-\mu}{1-\beta}\alpha + V_i^\alpha+\epsilon\,, \frac{1-\mu}{1-\beta}\alpha + V_k^\alpha+\epsilon'\}\,]\Big)\Bigg]$$

$$\Leftrightarrow \frac{1-\mu}{1-\beta}\alpha + V_i^\alpha = v_i + (1-\mu)\alpha + \beta\frac{1-\mu}{1-\beta}\alpha + \beta\Bigg[(1-\lambda\rho(\alpha))\times V_i^\alpha \lambda\rho(\alpha)\sum_k f\,\mathbb{E}_{\epsilon\epsilon'}[\,\max\{V_i^\alpha+\epsilon\,, V_k^\alpha+\epsilon'\}\,]\Big)\Bigg]$$

$$\Leftrightarrow V_i^\alpha = v_i + \beta\Big[(1-\lambda\rho(\alpha)) \times V_i^\alpha + \lambda\rho(\alpha)\ \eta_i^\alpha\ \Big]$$

As with $\eta_i(\alpha)$ we have from properties of the EV distribution:

$$\eta_i^\alpha = \sum_{k\in\Omega\setminus i} f \log\Big(\exp(V_i^\alpha) + \exp(V_k^\alpha)\Big)$$

Let me drop the α and use $\rho = \rho(\alpha)$ since this proof is conditional on α and we can rewrite it for any α such that $\rho(\alpha) > 0$. To show that all rankings imply the same preference ordering of firms, we rearrange the Mobility Value Function (MVF):

$$v_i = V_i\Big[1-\beta+\beta\lambda\rho\Big] - \beta\lambda\rho \times\ \eta_i \tag{7}$$

Let me denote by T the contraction mapping that correspond to this problem:

$$T_i(V) = v_i + \beta\Big((1-\lambda\rho)V_i + \lambda\rho \underbrace{\sum_{k\in\Omega\setminus i} f \log\Big(\exp(V_i) + \exp(V_k)\Big)}_{\eta_i(V)}\Big)$$

Where V is the vector of values and $T_i(V)$ is element i of $T(V)$. To show that T is indeed a contraction mapping, I use Blackwell's Theorem. Starting with monotonicity, let me consider vectors of values V and V' such that $V_i' \geq V_i \quad \forall i$. Note that by totally differentiating the option value, one gets $d\eta(V) = dV$. Hence, the option value implied by V' is weakly greater than for V :

$$\eta(V) = \sum_{k\in\Omega\setminus i} f \log\Big(\exp(V_i) + \exp(V_k)\Big) \leq \eta(V')$$

Furthermore, $1-\lambda\rho > 0$ and $\lambda\rho > 0$. Hence, monotonicity is verified :

$$T_i(V) \leq T_i(V') \quad \forall i$$

To show discounting, let me consider for some constant $a > 0$:

$$\begin{aligned}
T_i(V+a) &= v_i + \beta\Big((1-\lambda\rho)(V_i+a) + \lambda\rho \sum_{k\in\Omega\setminus i} f \log\Big(\exp(V_i+a) + \exp(V_k+a)\Big)\Big)\\
&= v_i + \beta\Big((1-\lambda\rho)(V_i+a) + \lambda\rho \sum_{k\in\Omega\setminus i} f \log\Big(\exp(V_i) + \exp(V_k)\Big) + \lambda\rho(K-1)\ f\ a\Big)\\
&= T_i(V) + \beta a
\end{aligned}$$

Since we have that $\beta < 1$, T is indeed a contraction mapping. Therefore, the Banach Fixed point Theorem implies that V is the unique solution to (7).

To show that the order of preferences according to v is similar as the one with V, let me assume it is not. For a pair of firm flow values such that $v_i < v_k$, I assume that $V_i > V_k$. From the proof above, there exists $V_i = T_i(V)$ and $V_k = T_k(V)$. I take their difference an obtain :

$$\begin{aligned}
&V_i - V_k = T_i(V) - T_k(V)\\
\Leftrightarrow &\Delta V_{ik} = \Delta v_{ik} + \beta\Big((1-\lambda\rho)\Delta V_{ik} + \lambda\rho\Delta\eta_{ik}(V)\Big)\\
\Leftrightarrow &\underbrace{\Delta V_{ik}(1-\beta(1-\lambda\rho)) - \beta\lambda\rho\Delta\eta_{ik}(V)}_{<0?} = \underbrace{\Delta v_{ik}}_{>0}
\end{aligned}$$

Note that (shown more in detail in the next proof) :

$$\Delta\eta_{ik}(V) = \sum_{j\in\Omega\setminus\{i,j\}} f\,\log\Big(\frac{\exp(V_i)+\exp(V_j)}{\exp(V_k)+\exp(V_j)}\Big)$$

Which gives that :

$$\mid \Delta\eta_{ik}(V) \mid \;\leq\; \mid \Delta V_{ik} \mid$$

As a result, I write $\Delta\eta_{ik}(V) = \chi\Delta V_{ik}$ for $\chi \in [0,1]$ and rewrite the previous inequality as :

$$\underbrace{\Delta V_{ik}(1-\beta(1-\lambda\rho(1-\chi)))}_{<0?} = \underbrace{\Delta v_{ik}}_{>0}$$

For all values of χ, we have that $1 > 1-\beta(1-\lambda\rho(1-\chi)) > 0$ which contradicts our assumption. Hence, any 2-elements preference relation on v is maintained with V. Similarly $V_i(\alpha)$ also conserves the preference order for all α because it is equal to V_i^α plus some constant.

***Proposition* 2:** *Under the same set of assumptions as Proposition 1, we have the following identity $\forall\alpha, i, j$:*

$$\Delta V_{ij}^\alpha \times h(\alpha,i,j) = \Delta v_{ij}$$

with $h(\alpha,i,j) = 1-\beta+\ \beta\,\lambda\,\rho(\alpha)\ (1-\chi_{ij}^\alpha)$ and χ_{ij}^α the difference of option values η_i^α ranging from 0 when (i,j) are the worst firms to 1 when they are the best ones.

As before, I keep notation light by dropping αs and write the difference operator as $\Delta V_{ij} = V_i - V_j$. Let us express the difference in Mobility Value Functions (MVF):

$$V_i - V_j = v_i - v_j + \beta(1-\lambda\rho)(V_i - V_j) + \beta\lambda\rho \times\ (\eta_i - \eta_j)$$
$$\Leftrightarrow \Delta V_{ij} = \Delta v_{ij} + \beta(1-\lambda\rho)\Delta V_{ij} + \beta\lambda\rho \times\ \Delta\eta_{ij}$$
$$\Leftrightarrow \Delta V_{ij}\Big(1-\beta+\beta\lambda\rho(1-\frac{\Delta\eta_{ij}}{\Delta V_{ij}})\Big) = \Delta v_{ij}$$

We can further detail $\Delta\eta_{ij}$:

$$\begin{aligned}\Delta\eta_{ij} &= \sum_{k\in\Omega\setminus i} f\,\log\Big(\exp(V_i)+\exp(V_k)\Big) - \sum_{k\in\Omega\setminus j} f\,\log\Big(\exp(V_j)+\exp(V_k)\Big)\\ &= f\,\log\Big(\frac{\exp(V_i)+\exp(V_j)}{\exp(V_j)+\exp(V_i)}\Big) + \sum_{k\in\Omega\setminus\{i,j\}} f\,\log\Big(\frac{\exp(V_i)+\exp(V_k)}{\exp(V_j)+\exp(V_k)}\Big)\\ &= \sum_{k\in\Omega\setminus\{i,j\}} f\,\log\Big(\frac{\exp(V_i)+\exp(V_k)}{\exp(V_j)+\exp(V_k)}\Big)\end{aligned}$$

Conditional on a set of values $\{V\}$, $\Delta\eta_{ij}$ is increasing in V_i and decreasing in V_j. Note that as V_i and V_j get high in the ranking, the terms V_k become relatively smaller and therefore in the limit $\Delta\eta_{ij}$ tends to ΔV_{ij} when all other V_k are negligible (relative to V_i and V_j). On the contrary, note that as V_i and V_j become negligible relative to V_ks - when they both are low in the ranking -, $\Delta\eta_{ij}$ tends to 0. We define χ as:

$$\chi_{ij} = \frac{\Delta\eta_{ij}}{\Delta V_{ij}}$$

As exposed in the *Proposition 2*, this term takes values between 0 and 1. It tends to 1 as firms are highest in the ranking and to 0 as firms get worst. The intuition is that as workers get in better firms, their option value η_i contains more of their own firm value and less of all other firms. Hence switching firms high in the ranking makes a difference in expected utility of the order ΔV_{ij}. Conversely, switching firms at the bottom makes no difference in terms of option value because they are both low (basically $\Delta\eta_{ij} \approx \mathbb{E}[V] - \mathbb{E}[V] = 0$). Note that χ_{ij} is increasing in the average of the ranks of i and j.

Finally, we get our expression for h, the function that maps expected value of firms to flow values:

$$\Delta V_{ij} \underbrace{\Big(1 - \beta + \beta\lambda\rho \times (1 - \chi_{ij})\Big)}_{h(i,j)} = \Delta v_{ij}$$

It is straightforward to generalize for all α.

I now show how I estimate this. To map v to V, one needs information on the entire distribution because it is needed to compute η. Without it, we cannot pin down the value functions:

$$v_i = V_i\Big[1 - \beta + \beta\lambda\rho\Big] - \beta\lambda\rho \times \ \eta_i$$

This could be solved as the fixed point of a system as hinted above. However, in practice, the estimation only requires v and h without needing to compute all V^α. Hence, I propose to use the function h and some approximation of χ to accelerate the estimation. I start by some definition. We call the Static Value Function (SVF), the utility derived from firm i when mobility is impossible.

$$V_i^* = v_i + \beta V_i^* \tag{SVF}$$

Here mapping expected values and flow values is easy: $v_i = (1-\beta)V_i^*$. Note that since it removes mobility, it kills the option value and therefore simplifies the mapping. This ranking V^* also satisfies quite obviously the same order of preferences as v. I then compute the shadow option value η^*:

$$\eta_i^* = \sum_{k \in \Omega \setminus i} f \log\Big(\exp(V_i^*) + \exp(V_k^*)\Big)$$

This option values uses information on v but not on V^αs. Next, I use this vector of option values to compute any χ_{ij}^α as:

$$\forall \alpha \quad \chi_{ij}^\alpha \approx \frac{\Delta\eta_{ij}^*}{\Delta V_{ij}^*}$$

	$\alpha = -3$	$\alpha = 0$	$\alpha = 3$
$\mathrm{V}(V^{\alpha})/\mathrm{V}(V^{*})$	99%	69%	52%
(Interc.)	0.000	0.005	0.009
β	1.000	1.000	1.001
R^2	1.0	0.999	0.997

TABLE 9: Prediction of χ^{α} with $\Delta\eta^{*}_{ij}/\Delta V^{*}_{ij}$

To motivate this modelling, I show numerically that for any α value this approximation performs very well. The intuition is that χ_{ij} is informative about the relative standing of i and j in the overall ranking. Since this ranking is similar to the ranking with v, I can recover this information and use it to predict the weight of the option value difference in the h function for all types $(1 - \chi_{ij})$. All-in-all, since α is not determinant to the term χ^{α}, I can neglect it at a low cost.

Table 9 shows the results from my simulation[28]. Remark that the fit gets less good as α gets larger. This is expected since mobility plays a larger role in determining the dispersion of utility. The first row shows the ratio of the variance of V^{α} over V^{*} (which is maximized because workers cannot move). We see that as workers get more productive, they are more mobile so their expected utility is less dispersed between firms. The results then show the coefficients for a linear regression of χ on $\Delta\eta^{\alpha}_{ij}/\Delta V^{\alpha}_{ij}$ simulated values. The fit is almost perfect for all relevant values of α (even for higher values not shown here). Basically, since mobility is quite low, the option value does not vary too much across workers. What changes more is the actual access to firms determined by the acceptance function ρ.

***Proposition* 3:** *Within a certain parameters region or when αs are not too dispersed within firms we can use:*

$$\int \rho(\alpha) P(V_j^{\alpha} \succ V_i^{\alpha}) dF_i(\alpha) \approx \rho(\alpha_i) P(V_j^{\alpha_i} \succ V_i^{\alpha_i})$$

with $\alpha_k = \int \alpha \, dF_k(\alpha)$

I show in Figure 5 how the mobility probability varies with the worker type α. The approximation proposed in *Proposition* 3 suggests that we can approximate the firm-level mobility rate by its of the average worker. This approximation is exact in 2 situations: (1) if there is only one type of workers in the firm ($F_i(\alpha) = \delta_{\alpha_i}(\alpha)$) or (2) if all α are in some interval where the function is linear. From the Figure, we see that in any other cases, my approximation will over- or under-estimate the true firm mobility rate. This derives from the fact that the function is concave above 0 and convex below. Hence taking the average type will systematically underestimate the mobility from firms with $\alpha_i < 0$ and systematically overestimate the mobility from firms with $\alpha_i > 0$. Hence, as mentioned in the main text, the dispersion in values can be interpreted as an upper bound for the true dispersion.

Finally, I show the proof for my variance decomposition with the Utility-Ladder (UL) and

[28] Simulation uses parameter values as estimated in the baseline results. I draw 200 firm flow values distributed logistic $(0, 5)$ and compute the expected values V^{α} for each type of worker. I then estimate the true χ^{α}_{ij} and compare it with the estimates given by χ^{*}_{ij}

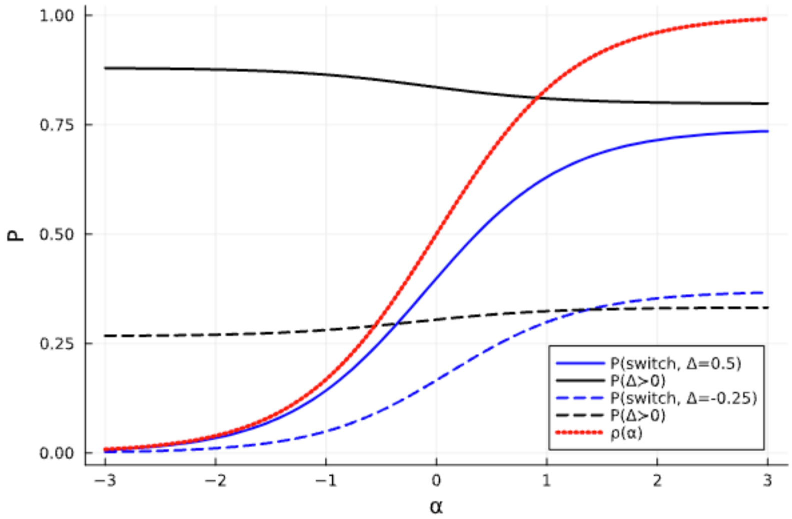

FIGURE 5: Mobility probability as a function of α decomposed into the dispersion-in-values and acceptance rate effects

Compensating Differentials (CD). First recall the formula for CDs:

$$\psi = \phi v + \tilde{CD} \quad \Leftrightarrow \quad \tilde{CD} = \psi - \phi v$$

where the CDs corresponds to the part of firm premia that is orthogonal to utility. It corresponds to the share of the variance that is actually compensated by wage differentials. We obtain the following decomposition:

$$v = \underbrace{\psi - \phi v}_{\tilde{CD}} + \underbrace{(1+\phi)v - \psi}_{UL}$$

We augment the utility dispersion to make CDs visible. The UL corresponds to the dispersion in utility plus the CDs. It can be though of as a counterfactual where firms do not compensate some amenity differentials with wages. We then compute the variance of those terms to arrive to the decomposition shown in the main text:

$$\begin{aligned} V(v) =& V(\tilde{CD} + UL) \\ =& V(\tilde{CD}) + V(UL) + 2Cov(\tilde{CD}, UL) \end{aligned}$$

$$V(\tilde{CD}) = V(\psi - \phi v) = V(\psi) + \phi^2 V(v) - 2\phi Cov(\psi, v)$$

$$\begin{aligned} Cov(\tilde{CD}, UL) &= Cov(\psi - \phi v, (1+\phi)v - \psi) \\ &= (1+\phi)Cov(\psi, v) - V(\psi) - \phi(1+\phi)V(v) + \phi Cov(\psi, v) \end{aligned}$$

$$\begin{aligned} 2Cov(\tilde{CD}, UL) + V(\tilde{CD}) &= -V(\psi) - \phi^2 V(v) + 2[(1+\phi)Cov(\psi, v) - \phi V(v)] \\ &= -V(\psi) - \phi^2 V(v) + 2\phi Cov(\psi, v) \\ &= -V(\tilde{CD}) \end{aligned}$$

Hence our result:

$$V(v) = V(UL) - V(\tilde{CD})$$

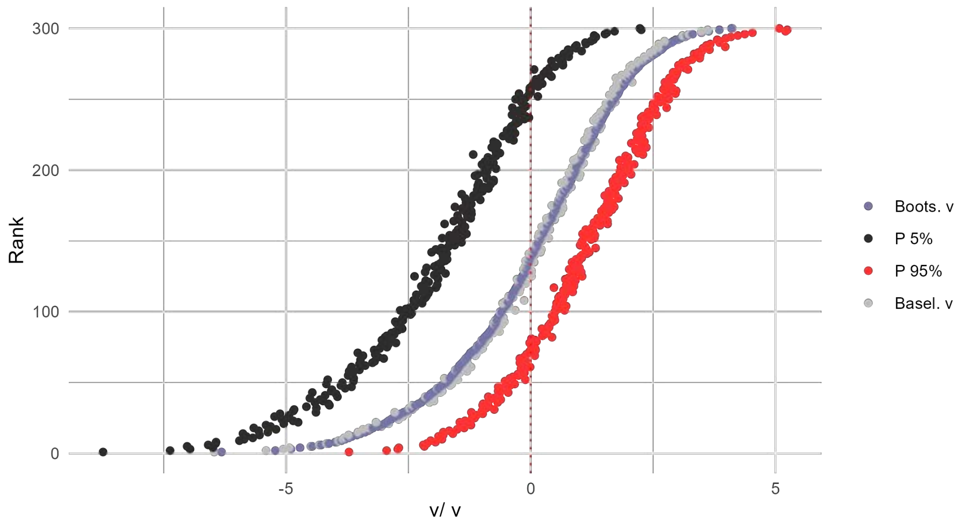

FIGURE 6: Binned mean and 90% interval of bootstrap estimates compared to Baseline estimates

B Robustness and Standard errors

In this Appendix, I discuss 3 aspects of my work that relate to robustness and uncertainty in my estimates. The first subsection presents the bootstrapped standard errors from the estimation. The second goes over additional estimation results that use more transitions to document the robustness of my estimates to those changes. Finally, I discuss the estimation of $1/\nu$ and the sensitivity of my results to any residual correlation between ψ and a.

1. Standard errors

Sorkin (2018) and Lehmann (2023) use bayes-empirical methods to shrink their estimates using the number of observations they have for each firm. Instead, I prefer to use a denser subset of firms that provides me more precise estimates. Indeed, I don't find my results to be overly dispersed when I compare them with estimates using more moves (see next subsection) but they might be imprecise. Hence, I perform a bootstrap of my estimates to evaluate the level of precision that I attain.

I bootstrap the main results 300 times by sampling without replacement in the original dataset. Figure 6 shows the 5^{th} and 95^{th} percentiles of my results for each firm value[29]. The Figure shows wide and left skewed confidence sets. The dashed line shows that some 13% of firms are significantly better than the average firm and 23% are significantly worse, the rest has the average falling in the 90% interval of estimates. Since most firms are actually close in the distribution, it is very hard to tell them apart. Still, it is possible to rank them and account for the uncertainty in the ranking. Note that at this stage, vs could be in any scale because scale does not matter to discuss the preferences, only to convert them to the log euro scale. I account for uncertainty by computing the ranking confidence sets as in Mogstad et al. (2020). Marginal and simultaneous sets are displayed in Appendix Figure 8. Assuming that estimates are normally distributed - which they are not, we get that individual rank (marginal) and the global ranking (simultaneous) are contained in these sets with probability 90%. Under this light, my estimates clearly do not allow to draw inference at the

[29]They are binned for anonymity concerns.

firm level, but they provide a robust description of the French economy. Keep in mind that I have as little as 5 moves to locate each firm in the ranking.

2. Sample selection

In Table 10, I provide a subset of results for different data specifications. I describe the samples and discuss the results in turn. The first row reports my baseline results for reference. To ease comparison, this table reports results for the set of firms contained in the baseline results. I estimate the values for the full sample and then only report statistics computed with the subset of firms that intersect with the baseline SCS.

- **Including ENE transitions:** As a first test, I keep the baseline sample and estimate values with all transitions between tenured jobs, including those through more than 30 days of unemployment. Since ENE transitions are more likely to be involontary, vs might be biased and seem to deliver systematically more utility to workers making an ENE move. This results into a lower $1/\nu$ because ENE moves are not towards higher paying firms, hence the regression spuriously identifies a high role for amenities. Similarly, this bias increases significantly the size of residual utility and therefore the contribution of amenities to total compensation dispersion.

- **Not filtering for firm size:** In this estimation, I use a larger sample of firms (~ 7000) to see how omitting firms not in the SCS might produce noisier estimates. In particular, I do not perform the third filtering step that removes half of the transitions by keeping firms that have 30 workers per period and 5 in- and out-moves. I find estimates very much in line with the baseline results (for the firms included in the baseline), if anything with more dispersion in utility.

- **Including CDD workers:** For this test, I keep all full-time workers, no matter their tenure. Bear in mind that in the main sample I filter for CDI-CDI transitions. Therefore, this new sample adds CDD-CDI, CDI-CDD and CDD-CDD transitions and might also suffer from containing workers that involontarily switch firms. This results in a lesser contribution of amenities compared to the previous results. This could reflect the fact that workers pre-tenure do not value amenities as much as tenured workers, motivating evidence for my restriction to tenured employees. Also, this could suggest some tenure effect whereby tenure compensates for amenities (as would wage CDs), this is left for future work. However, it could very well derive from the fact that more transitions allow for less noisy estimates and therefore a lower contribution of amenities to compensation dispersion - as suggested by the lower share of CDs which are basically residuals from $v \sim \psi$.

- **Longer time span:** I then repeat the whole estimation as in the baseline but including data for the years 2016-2018. My baseline only contains 5 years to garantee the stability of ψ and a over the period but it is interesting to use a longer period and still assume stability. The results are very close to the baseline in all regards. They also suggest that more data helps to estimate more precisely CDs but do not point towards qualitatively different results in terms of compensation dispersion. I restrict my analysis to the first period but could also compare how my estimates evolve with time. This is also left for future work.

Lastly, Sorkin (2018) shows by Monte Carlo simulations that his estimator does not quantitatively provide too dispersed estimates. This is consistent with my sample checks not

	$(1+\beta)/\nu$	$1/\nu$	$\tilde{\text{CD}}$	V(a)	2cov(a,ψ)	2cov$(a,\bar{\alpha})$
Baseline	10.8	7.1	47%	0.049	0.008	0.095
EE + ENE	8.9	4.2	47%	0.1010	0.0175	0.1559
No filter loop	11.8	7.2	46%	0.0540	0.0097	0.1044
CDI + CDD	7.5	7.0	38%	0.0290	0.0021	0.0610
2010-2018	10.9	6.9	36%	0.0358	0.0103	0.0784

Table 10: Results with broader samples

delivering great differences depending on the number of transitions. Indeed, row 5 shows less CDs but very close estimates.

3. Estimation and robustness to scale measurement

In this subsection, I motivate my choice of specification for estimating $1/\nu$ and provide some illustration on how relaxing this assumption would affect my results.

Table 11 provides all the steps towards finding a suitable estimate for $1/\nu$. First, recall that the regression $v \sim \psi$ might be biased depending on the correlation β between a and ψ. Of course the coefficient resulting from the regression using workers mobility choices $\Delta v \sim \Delta\psi$ might suffer from a similar bias because workers sample firms from an underlying distribution that has some unobserved correlation. Therefore, my estimate in row 2 might suffer from a similar bias as row 1 even though it is way lower. Now consider the fact that workers are originally located in some firm and that they have heterogeneous mobility rates. Hence my estimate from $\Delta v \sim \Delta\psi$ might not reflect the underlying distribution of (ψ, a) anymore but some function of the location of all workers on the job-ladder and their individual productivity. In addition, note that depending on the location in the ladder, β could change, for example at the top all firms could pay high wages but provide very different amenities, dampening the relationship between v and ψ. Consider for example in row 4, the regression using only moves conditional on workers staying in the same industry and the same occupation. Those workers could very well want to keep the same working conditions while moving to a higher-paying firm. On the contrary they might not care so much about the wage premium but about the working conditions given that they are likely to have some bargaining power from their experience. The very low estimates seem to suggest that the second narrative dominates in the data.

To sidestep these considerations, I make the assumption that workers that go through unemployment search jobs by prioritizing firm premia compared to total compensation. The intuition is that they might be more constrained financially and therefore disregard amenities. This might not completely reduce the fact that workers sample from a correlated distribution, but at least reduces the bias and indicate its direction. Comparing rows 2 and 3 suggests that amenities are correlated positively with firm premia (and conversely for a_0). I find that the results from a regression using ENE movers provide systematically consistent estimates with regard to the resulting correlation between a and ψ. In addition, I find that this relationship increases with income before the move, suggesting that higher-income non-employed workers might be waiting to find a better offer (pushing up the correlation $\Delta v \sim \Delta\psi$). Therefore, I only consider workers that are in the bottom 25^{th} percentile of the firm values. My preferred specification is given by the starred estimates. Using EE moves from the bottom of the distribution indeed shows that individuals down the ladder care a lot about wage premia,

Model	v	v_0
$v \sim \psi$	10.8	3.9
$\Delta v \sim \Delta\psi$ | EE	6.4	4.4
$\Delta v \sim \Delta\psi$ | ENE	6.7	3.8
$\Delta v \sim \Delta\psi$ | within	3.5	1.3
$\Delta v \sim \Delta\psi$ | EE $\cap$ 25$^{\text{th}}$	9.4	8.9
$\Delta v \sim \Delta\psi$ | ENE $\cap$ 25$^{\text{th}}$	7.1*	6.6*

TABLE 11: Taste shock scale measurement

but it doesn't mean that they don't consider the amenity which is still correlated with firm premia.

To address the concern that my assumption could be misguided, I provide a simple sensitivity analysis. If indeed there subsists some correlation $\tilde{\beta}$ after restricting my regression to the set of ENE movers at the bottom of the firm values distribution, I can still recompute the true $1/\nu$ as a function of some residual correlation $\tilde{\beta}$ and my estimate $\hat{\gamma}$ such that $1/\nu = (1+\tilde{\beta})/\hat{\gamma}$. I show the full utility decomposition as a function of $\tilde{\beta}$ in Appendix Figure 10 (LHS). Changing $1/\nu$ only affects my estimates of a. Given that I find a correlation between a and ψ of 0.24 (Table 12 and Appendix Figure 9), a residual correlation of around $-.3$ faced by ENE movers would overturn my result. Regarding sorting on amenities, this correlation would need to be as low as $-.8$, a very unrealistic number. On the contrary, a positive unobserved correlation - due to the underlying correlation in job offers to unemployed - would only increase the role of amenities in the utility dispersion. Given my estimate in the regression $v \sim \psi$ this is most likely direction of such potential bias.

C Additional graphs and tables

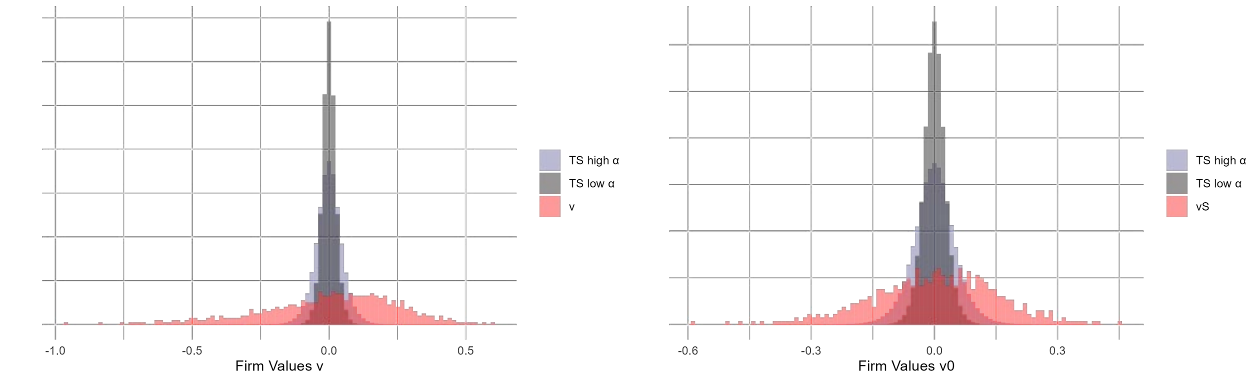

FIGURE 7: Taste shock relative to the estimated dispersion of firm flow values v (LHS) and v_0 (RHS)

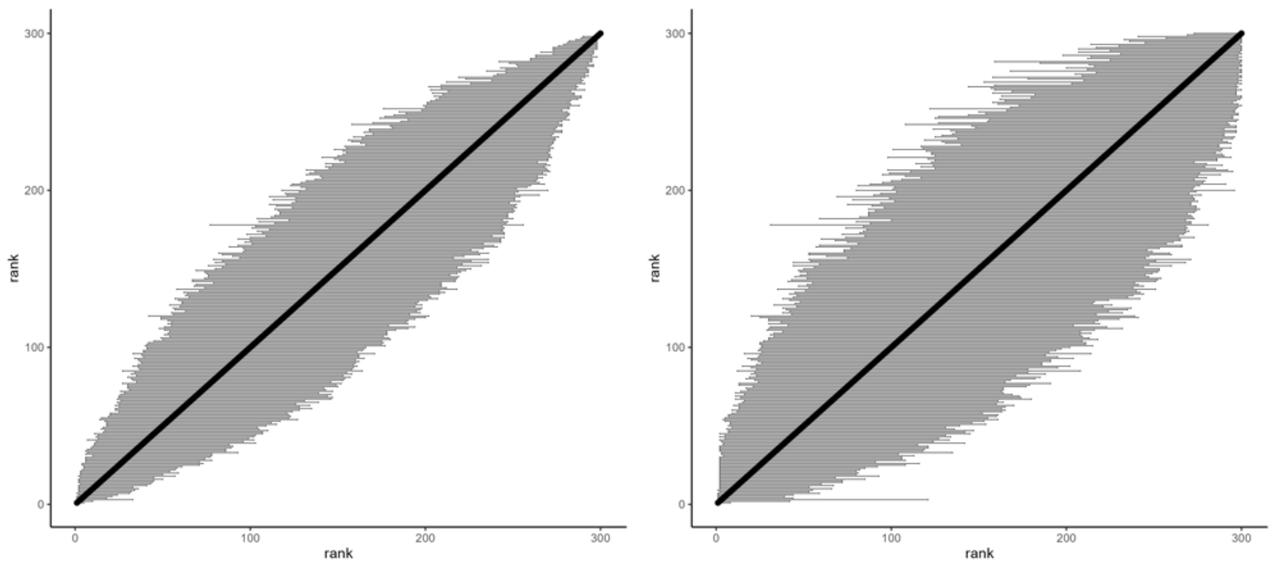

FIGURE 8: Marginal (LHS) and Simultaneous (RHS) confidence sets for the ranking of firms

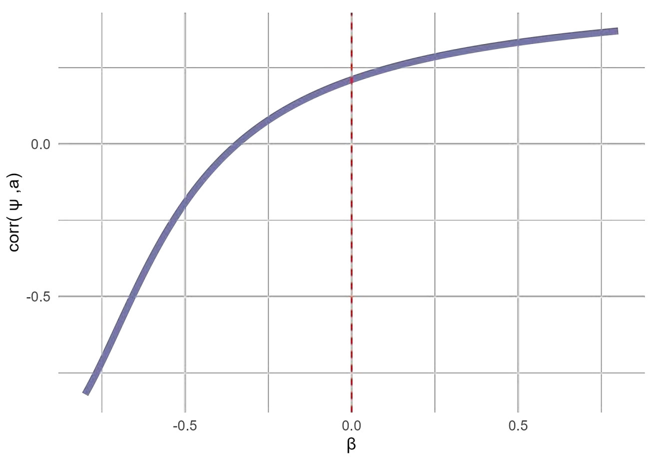

FIGURE 9: True Cor(ψ, a) as a function of the residual unobserved correlation $\tilde{\beta}$

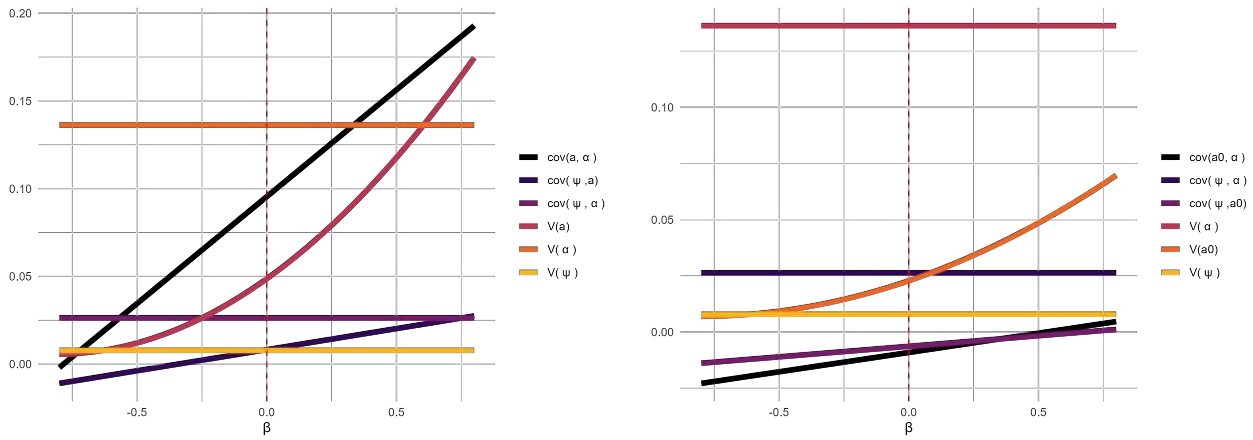

Figure 10: Sensitivity of the full utility decomposition as a function of the residual unobserved correlation $\tilde{\beta}$

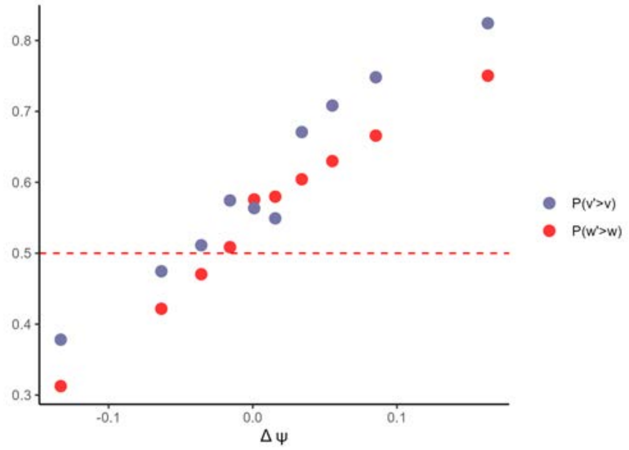

Figure 11: Probability of moving to a higher v firm and getting a higher wage conditional on firm FE difference $\Delta\psi$ upon moving

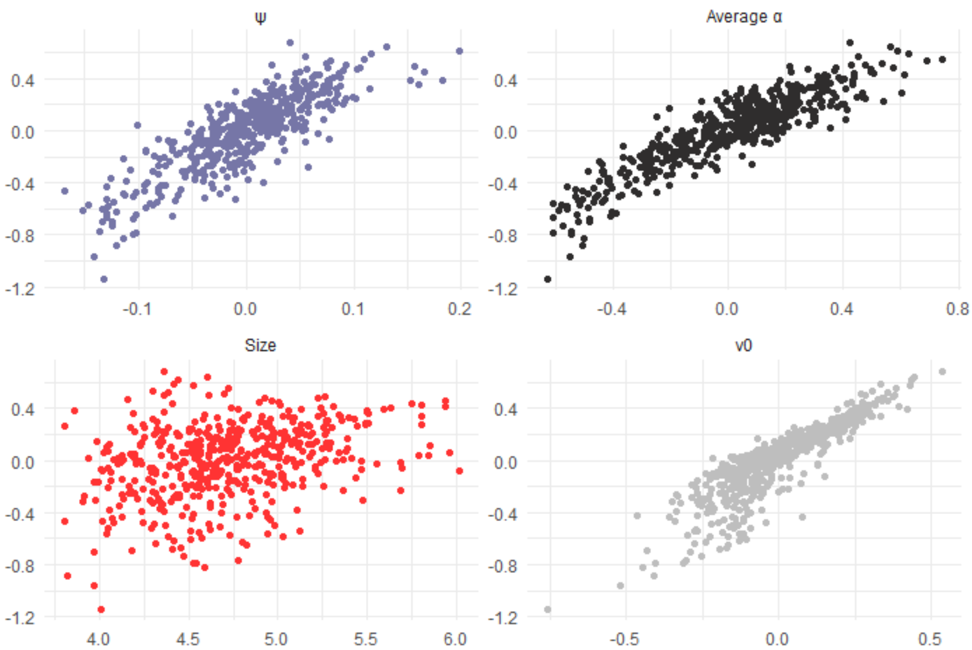

Figure 12: Panel of variables against v: ψ, $\bar{\alpha}$, firm size and v_0 (all binned)

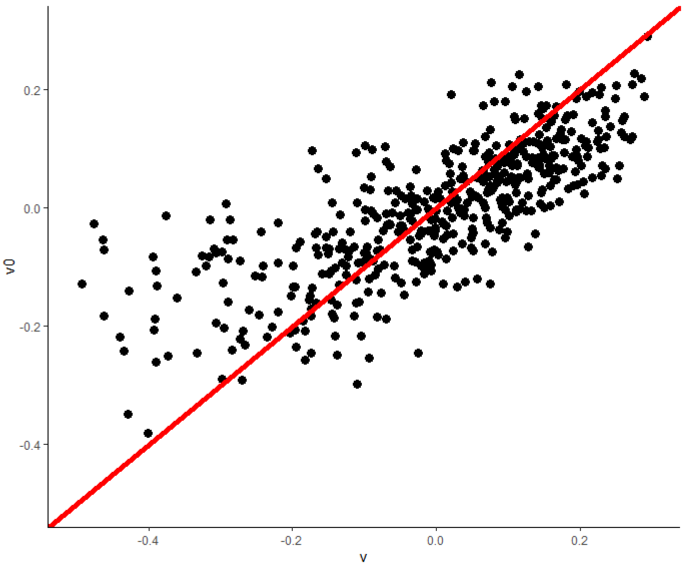

Figure 13: Binned scatterplot of v_0 and v

	v	$\bar{\alpha}$	ψ	a	v_0	a_0
v	-	0.65	0.53	0.94	0.66	0.37
$\bar{\alpha}$		-	0.40	0.59	-0.08	0.15
ψ			-	0.21	0.34	-0.24
a				-	0.62	0.52
v_0					-	0.82
a_0						-

Table 12: Correlation of estimates

	ψ	a	ψ	a	ψ	a
Intercept	-0.317***	-0.621***	0.128	0.484	0.096	0.431
	(0.087)	(0.188)	(0.330)	(0.580)	(0.313)	(0.492)
Weekly hours / 100	0.498***	1.692***	-1.586	-4.393	-1.299	-3.734
	(0.214)	(0.576)	(1.704)	(3.315)	(1.634)	(2.863)
$(\text{Hours}/100)^2$			3.360	8.338*	2.815	6.942
			(2.260)	(4.732)	(2.184)	(4.085)
a or ψ					0.065***	0.416**
					(0.025)	(0.151)
SE sector cluster	Y	Y	Y	Y	Y	Y
R^2	0.07	0.04	0.07	0.05	0.10	0.07

Table 13: Regression of firm premia and amenities on average weekly hours

Sector	A16	v rank	v_0 rank
Activities auxiliary to financial services and insurance	KZ	1	9
Leather and footwear industry	C	2	1
Insurance	KZ	3	2
Financial services, except insurance and pension funding	KZ	4	8
Manufacture of other transport equipment	C	5	4
Other manufacturing industries	C	6	3
Programming and broadcasting	JZ	7	13
Chemical industry	C	8	5
Publishing industry	JZ	9	28
Beverages manufacturing	C	10	17
Pharmaceutical industry	C	11	21
Scientific research and development	MN	12	26
Manufacture of machinery and equipment	C	13	20
Manuf. of fabricated metal products, exc. mach. and equip.	C	14	27
Manufacture of computer, electronic and optical products	C	15	32
Other specialized, scientific and technical activities	MN	16	22
Wholesale trade, except of motor vehicles and motorcycles	GZ	17	19
Head office activities; management consulting	MN	18	31
Manufacture of rubber and plastic products	C	19	34
Air transport	HZ	20	25
Manufacture of other non-metallic mineral products	C	21	36
Real estate activities	LZ	22	6
Automotive industry	C	23	44
Electrical equipment manufacturing	C	24	41
Information services	JZ	25	42
Telecommunications	JZ	26	48
Repair and installation of machinery and equipment	C	27	18
Water collection, treatment and distribution	DE	28	24
Postal and courier activities	HZ	29	11
Motion picture, video and television program production	JZ	30	49
Sports, leisure and recreational activities	RU	31	43
Building construction	FZ	32	12
Advertising and market research	MN	33	52
Teaching	OQ	34	51
Food industry	C	35	38
Rental and leasing activities	MN	36	30
Programming, consulting and other IT activities	JZ	37	55

Table 14: Ranking on v and v_0 for the granular sectoral decomposition

Sector	A16	v rank	v_0 rank
Administrative and other business support activities	MN	38	37
Civil Engineering	FZ	39	16
Metallurgy	C	40	56
Electricity, gas, steam and air conditioning prod. and distr.	DE	41	15
Clothing industry	C	42	7
Architectural and engineering & technical control and analysis	MN	43	52
Legal and accounting activities	MN	44	57
Travel agencies, tour operators, reservation services and related	MN	45	50
Warehousing and auxiliary transport services	HZ	46	35
Specialized construction work	FZ	47	23
Wastewater collection and treatment	DE	48	45
Waste collection, treatment and disposal & recovery	DE	49	14
Land and pipeline transport	HZ	50	10
Retail trade, except of motor vehicles and motorcycles	GZ	51	46
Sale and repair of automobiles and motorcycles	GZ	52	29
Catering	IZ	53	33
Hospitality services	IZ	54	47
Activities for human health	OQ	55	40
Building services and landscaping	MN	56	39
Investigation and security	MN	57	54

Table 15: Ranking on v and v_0 for the granular sectoral decomposition (Continued)

Sector	code
Mining, energy, water, waste management and remediation	DE
Manufacturing	C
Construction	FZ
Trade	GZ
Transportation and storage	HZ
Accommodation and catering	IZ
Information and communication	JZ
Financial and insurance activities	KZ
Real estate activities	LZ
Scientific and technical activities & admin. and support services	MN
Public administration, education, health and social work	OQ
Other service activities	RU

Table 16: Sector codes