Individual Wage Dynamics, Continuous Firm Types and Path Dependence

Richard Blundell   Thomas Breda   Marc Chan   Jean-Marc Robin*

March 21, 2025

Abstract

We use French matched employer-employee data to estimate a structural model of wage dynamics. Adding further frictions to the sequential auction model, we find that job-to-job mobility is predominantly inefficient. That is moves are not purely in response to a financial surplus. Allowing for between job-to-job moves that should have occurred (higher surplus with an external firm) but did not, and moves that did occur but should not have (lower surplus in destination firm), we estimate that the proportion of employees whose last state change is inefficient is about 35%. We also find that there is strong worker-firm sorting, but this does not show up in wages. We investigate in detail the role of the dynamics of match-specific wage shocks and examine the decomposition of the variance of wages and of the match surplus. We find that the match output, reflecting firm effects and worker-firm complementarities, is small while the match-specific effect dominates.

*Blundell, UCL, IFS, r.blundell@ucl.ac.uk; Breda, PSE, CNRS; Chan, University of Melbourne; Robin, Sciences Po. We thank , and seminar participants at , for helpful comments and suggestions. Blundell thanks the ESRC for support under the Centre for the Microeconomic Analysis of Public Policy (CPP) at IFS, grant number ES/M010147/1. Robin acknowledges support from the European Research Council (grant reference ERC-2020-ADG-101018130).

1 Introduction

Structural search models provide an elegant framework for the analysis of job-to-job transitions and wage dynamics. For example, the sequential auction model of Postel-Vinay and Robin (2002), which we use to cast our analysis, is a well developed model for individual wage dynamics and employment mobility that allows for both worker and firm heterogeneity, transitions in and out of employment, and transitions between employers. In this equilibrium search model, wages and mobility are determined by Bertrand competition between competing employers. If the incumbent firm wins the auction the wage may rise. If the "poaching" firm wins, the worker moves, generally also with a wage rise. However, despite many extensions,[1] the model implies restrictions on the path dependence of wages that lead to empirical predictions that are at odds with some observed patterns of wage dynamics and job-to-job mobility. In particular, the model does not allow for substantial wage cuts when employees move between jobs. However, as Jolivet, Postel-Vinay, and Robin (2006) and Tjaden and Wellschmied (2014) note, job-to-job transitions of employees are often observed with significant wage cuts. There is also considerable empirical evidence that mobility costs can deter workers from accepting jobs with better pay (see Cruz, Milet, Olarreaga, and Solleder, 2024, for a recent reference), although mid-career change can lead to higher employment at older age (OECD, 2024). We also find such wage cuts in our analysis of French administrative data, and the path of transitions with firm tenure suggest that some workers remain with an employer even though it appears that better offers are available elsewhere.

By comparison, statistical models of panel data earnings dynamics impose few prior restrictions on the dynamics. Most models are flexible dynamic factor models with a worker fixed effect and linear permanent and transitory components (see Moffitt and Gottschalk, 2002, for example). Some specifications assume non linear dynamics. Meghir and Pistaferri (2004) is an early example featuring ARCH innovation errors. Recent specifications allow for more general nonlinear persistence, pointing to asymmetric shocks where the sign and size of earnings shocks are associated with different persistence (see Arellano, Blundell, and Bonhomme, 2017). These models effectively allow a reset in path dependency of the dynamic earnings process when unusual shocks occur. However, they are not able to assess the structural determinants of firm-to-firm transitions and wage dynamics.[2] In particular, they are

[1] More flexible rents (Dey and Flinn, 2005, Cahuc, Postel-Vinay, and Robin, 2006); richer wage dynamics (piece rate wage bargaining in Bagger, Fontaine, Postel-Vinay, and Robin, 2014; productivity shocks in Lise, Meghir, and Robin, 2016); sorting (Lise, Meghir, and Robin, 2016, Lise and Robin, 2017, Bagger and Lentz, 2019, Lamadon, Lise, Meghir, and Robin, 2024); large firms (Bilal, Engbom, Mongey, and Violante, 2022); match-specific heterogeneity (Taber and Vejlin, 2020); multi-dimensional heterogeneity (Lise and Postel-Vinay, 2020, Lindenlaub and Postel-Vinay, 2023); amenities and mobility costs (Lamadon, Lise, Meghir, and Robin, 2024).

[2] Typically nonlinear panel data models do not distinguish firm-to-firm moves (quits or displacements),

silent on the extent to which job-to-job moves are efficient. That is whether moves occur in response to financial surplus.

The usual way of explaining that many job-to-job moves are not responding purely to financial surplus is to incorporate extra adjustment costs into the transition model (see e.g. Lamadon, Lise, Meghir, and Robin, 2024, Sorkin, 2018, and Lentz, Piyapromdee, and Robin, 2023). These transaction costs involve a reset in the path dependency of earnings which is empirically useful.[3] The problem with amenities or mobility costs, as with measurement errors, is that they are not directly observables (at least not in totality). A large number of papers thus use indirect inference or simulated moment estimation to estimate structural models with many latent variables. But indirect inference has the particular disadvantage of having to choose a tractable auxiliary model. This usually involves some arbitrariness and it is generally difficult, if not impossible, to show that the mapping from structural to auxiliary parameters is injective.

Hagedorn, Law, and Manovskii (2017) were the first to study and prove identification of a search-matching model with no search on the job from auxiliary parameters. Recently, Lamadon, Lise, Meghir, and Robin (2024) use the finite mixture model of Bonhomme, Lamadon, and Manresa (2019) to construct the auxiliary model, and they show identification of a sophisticated sequential auction model with discrete types and a parametric distribution of mobility costs/amenities by the reduced form. The finite mixture is however a rather coarse approximation, with a small number of worker and firm types, of what must be the true heterogeneity distributions.

Our first contribution in this paper is to propose a tractable equilibrium framework for individual wage dynamics and employment mobility that can explain these key dynamic features of the data while providing sufficient structure to explore the mechanisms underlying firm and worker behavior. Specifically, we shall assume an inefficient sequential auction mechanism where the auction can be won by the second bidder with some positive probability (to be estimated). Whether a worker moves to a dominated poacher — we call this a displacement — or stays with a dominated incumbent may reflect the presence of amenities or mobility costs that we do not model directly. This innovation will allow us to use maximum likelihood for estimation and to model worker and firm heterogeneity as continuous

even though there is evidence that persistent shocks load heavily on firm-to-firm moves (see Altonji, Smith Jr., and Vidangos, 2013). They also typically do not examine entrants or exits in any systematic way. Important exceptions are Altonji, Smith Jr., and Vidangos (2013) and Taber and Vejlin (2020) that allow for a comprehensive set of moves; the latter using a Roy model to show the importance of non-pecuniary aspects of jobs, finding that one-third of all choices between jobs would have resulted in a different outcome if the worker only cared about wages.

[3]See Di Addario, Kline, Saggio, and Søvsten (2023) for recent evidence that mobility wages are less dependent on past firm type than predicted by the sequential auction model.

distributions.

A second feature we add to the search-matching framework is one that has been understudied. The widespread availability of matched employer-employee data in the last 25 years, has rendered possible the estimation of both worker and employer effects on wages.[4] The contribution of worker heterogeneity to the total wage variance is usually found to be large (from 50% up to 75%). The firm contribution is found to be small, as is that of the covariance between worker and firm effects. Match-specific shocks to productivity, when estimated (rarely), are comparable in magnitude to the firm effects.[5] However, this regression framework tells us nothing about firm-to-firm dynamics. Recently, Bonhomme, Lamadon, and Manresa (2019) and Abowd, McKinney, and Schmutte (2019) have proposed a dynamic model of both wages and job mobility with discrete worker and firm heterogeneity. They confirm the previous results and also find evidence of a partial disconnection between job preferences and wages.[6] Overall, a new understanding emerges that sorting is mainly driven either by exogenous market segmentation, or by additional frictions such as mobility costs and amenities that shift job preferences away from pure financial rationality.

In this paper, we also assume that workers sample firm types conditional on their own type (market segmentation). Moreover, anticipating that the match surplus determining job preference may not be the product of worker and firm type complementarities, we add to it an idiosyncratic random match effect. We discuss why we believe that we can separately identify market segmentation, Beckerian complementarities, an idiosyncratic match effect and an inefficient sequential auction mechanism.

That is a lot to put into one model. We have therefore decided to give up on something that we consider to be less essential, namely forward-looking behavior. The model in this paper is a static approximation to the true forward-looking model. In the absence of investment decisions, this should not be a major limitation. We will lose the prediction of the Postel-Vinay-Robin model that workers are willing to accept a lower wage in a good match

[4]See Abowd, Kramarz, and Margolis (1999), Abowd, Creecy, and Kramarz (2002), Andrews, Gill, Schank, and Upward (2008), Gruetter and Lalive (2009), Iranzo, Schivardi, and Tosetti (2008), Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer, and Woodcock (2009), Card, Heining, and Kline (2013), Woodcock (2015), de Melo (2018), Song, Price, Guvenen, Bloom, and von Wachter (2018), Kline, Saggio, and Solvsten (2020), Babet, Godechot, and Palladino (2022), Azkarate-Askasua and Zerecero (2023). The fixed effects are estimated using OLS. Second-order bias corrections were developed by Andrews, Gill, Schank, and Upward (2008), Kline, Saggio, and Solvsten (2020) and Azkarate-Askasua and Zerecero (2023).

[5]See Sørensen and Vejlin (2011), Sørensen and Vejlin (2013), Abowd, McKinney, and Schmutte (2019), Taber and Vejlin (2020) and Jinkins and Morin (2018). Sørensen and Vejlin (2011) and Jinkins and Morin (2018) measure the presence of worker and firm effects in an equation for worker wage growth, hence implying match effects in level wages.

[6]See also Sorkin (2018) and Lentz, Piyapromdee, and Robin (2023). The latter build on Bonhomme et al. by adding structure to the employment transition probabilities. They separate the chance of receiving an offer from a specific firm type from the choice of accepting it.

today in exchange of better future prospects. But this effect has not been found sufficient to explains the amount of wage cuts experienced by workers when they move between firms.

Under these assumptions, and further parametric restrictions, we show that a worker's log wage is the sum of a reservation log wage and a worker surplus. The reservation log wage is the sum of an experience effect — we allow workers to be born, age and retire — and worker innate ability (worker heterogeneity). The match surplus is the sum of a function of worker ability and firm technology (firm heterogeneity) and a match-specific component. The worker surplus is a fraction of the match surplus that evolves stochastically over time as a result of match draws.

As we have already mentioned, we will model unobserved worker and firm heterogeneity as two continuous latent variables, which is consistent with theoretical search-matching models. We will also use Maximum Likelihood, as in Bonhomme, Lamadon, and Manresa (2019) and Lentz, Piyapromdee, and Robin (2023). Bonhomme et al. develop a two-stage estimation procedure. In the first stage, they classify firms using k-means clustering based on firm-level wage distributions. In the second stage, they maximize the observed worker log-likelihood given firm clusters using the EM algorithm. Lentz et al. iteratively classify firms based on the expected log-likelihood after convergence of the worker EM stage. In this paper, we do something similar. In a way that is reminiscent of the Krusell-Smith model, we first posit that the latent firm type is a linear index of some firm-level variables (say the average and the standard deviation of the wages paid by the firm). Assuming that the latent firm type is observed up to a small number of factor loadings is a considerable reduction of dimensionality. Then, conditional on the firm index, we maximize the expected likelihood using numerical approximations of the integrals. The parameters of the firm index are estimated in an outer loop. The very tight risk structure of the model never requires integrating out more than two latent variables, the worker type and the current worker surplus for the first wage observation.

In order to average the log-likelihood over these latent variables, we need first to calculate the equilibrium joint distribution of the worker's age and ability, the match surplus and the worker surplus. This four-dimensional distribution solves a system of differential equations that is tedious to solve, but once this is done, the likelihood is rather straightforward to calculate and to maximize. Finally, there is one last econometric innovation. We observe wage cuts within employment spells and wage increases may or not reflect employer competition. We therefore develop a noise filtering algorithm on the raw data to remove wage cuts and to keep only wage increases greater than some arbitrary threshold (5%). This noise reduction procedure only cuts 2% of the total wage variance. We view this as a small cost compared to incorporated measurement error in the likelihood.

Applying this framework to the French matched employer-employee data we find that a large fraction of job-to-job moves (35%) are inefficient. That is those cannot be explained by financial surplus. In addition, we uncover a strong role for worker-firm sorting but one that does not show up in wages and does not result from correlated worker and firm effects. Finally, idiosyncratic firm effects contribute little to wages and job preferences compared to match-specific effects. Decomposing the variance of wages and of the match surplus, we find that the match production function reflecting worker-firm complementarities is small. In contrast we find that the match-specific effect plays an important role. We provide a detailed investigation of the role of the dynamics of match-specific wage shocks.

The rest of the paper is organized as follows. In Section 2 we present the model of individual wage dynamics. In Section 3 we discuss the steady-state economy. Section 4 presents the data and Section 5 the estimation procedure. Section 6 discusses the results. Section 7 concludes.

2 An inefficient sequential auction model

2.1 Human capital and match output

We consider a population of infinitely-lived firms and short-lived workers. Time is continuous. A worker has human capital $R = \exp(\gamma(t) + x)$, where $\gamma(t)$ is a deterministic function of age/potential experience t (with $\gamma(0) = 0$) and x is unobserved ability. Human capital is understood as a worker's productivity in non-wage employment, or simply non-employment (where we include self-employed as well as unemployed in this definition throughout), and is therefore the minimum wage he or she will accept from any employer.

Firms are characterized by a technological index y. The output of a firm-worker match is Re^z, with $z = f(x, y) + v$, where $f(x, y)$ is the contribution to match output of worker and firm characteristics, and v is a match-specific component. The match output is the maximum wage that the firm is willing to pay the worker. We call z the match surplus and only matches with $z \geq 0$ are viable (preferred to non-employment).

We assume that x, y and v are continuous variables. Both x and y are fixed characteristics. If a worker returns to the same firm after an interruption, she sees the same y. But v is a random effect. If a worker returns to the same firm after an interruption, she draws a new v.

2.2 Wages out of non-employment

Non-employed workers receive job offers at Poisson rate λ_0. A worker (t, x) draws an employment offer (y, v). The match is formed if surplus z is nonnegative. In which case, the worker accepts the offer and receives a wage inside the bargaining set $[R, Re^z]$. Specifically, we assume that non-employed workers exit to employment with a wage W (or log wage w) such that

$$w := \ln W = \gamma(t) + x + \beta z. \tag{1}$$

The job contract specifies a wage indexed on experience. It remains fixed until some alternative offer forces the incumbent employer to renegotiate. From now on, unless specified otherwise, all wages are in logs and denoted using lower-case letters.

2.3 Wages on the job

Wage renegotiation and job mobility are triggered by concurrent job offers ("poaching"). In the data, it is impossible to observe productivity shocks or workers' learning on the job. But we do see employees changing employers. It is likely that workers are infrequently in situations where they have the possibility to change employer. Sometimes they do, sometimes they don't. It is this imperfect competition among employers for workers that gives market power to firms that can pay their workers less than the marginal productivity of their jobs. We distinguish three different types of poaching environments.

Normal negotiation. The normal poaching game is a static version of the wage renegotiation processes in Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006). Bertrand competition forces the two firms to bid their reservation values equal to the match log value $\gamma(t)+x+z$. The first bidder wins and the worker gets an extra share of the surplus by Nash bargaining with the second bid as threat point.

Specifically, with probability λ_1, an employee draws an alternative offer z' yielding the following outcome:

- (stay) if $z' \leq z$, the worker keeps job z and the worker's wage increases to

$$w' = \max\left\{w, \gamma(t) + x + \beta z + (1-\beta)z'\right\};$$

- (move) if $z' > z$, the worker moves to z' with a new wage

$$w' = \gamma(t) + x + \beta z' + (1-\beta)z.$$

Note that

$$\gamma(t) + x + \beta z + (1-\beta)z' = \beta\left[\gamma(t) + x + z\right] + (1-\beta)\left[\gamma(t) + x + z'\right]$$

where $\gamma(t)+x+z$ is the match output or productivity. Therefore, the hiring wage is a simple linear combination of the pre- and post-transition match productivities.

We now consider two additional frictions that increase the chance of narrowing the wage-productivity gap.

Worker moves anyway. Many job-to-job transitions of employees are observed with wage cuts. We therefore introduce the possibility that the auction is played, firms bid their reservation values (i.e. match productivity), but the worker quits anyway, accepting whatever the poacher has offered irrespective of the incumbent employer's bid. Perhaps the worker wanted to quit because she was fed up with her job, or perhaps the new job comes with amenities that make up for some wage loss. We assume that competing employers do not factor in the risk that the worker may not accept the best financial offer.

Specifically, with probability λ_2 a worker employed with match surplus z draws an alternative job offer with surplus z' resulting in the following outcome:

- she leaves employment if $z' < 0$;
- if $z' \geq 0$, she takes the new job with a wage

$$w' = \gamma(t) + x + \min\{\beta z' + (1-\beta)z, z'\}.$$

  If $0 \leq z' \leq z$, the worker moves to the poaching firm although she should not (displacement) and Bertrand competition gives the whole surplus to the worker. If $z' > z$, this is like normal poaching.

The worker becomes non-employed with probability λ_2 times the probability of drawing a non viable job with surplus $z' < 0$. This endogenous layoff risk depends on the worker's ability as more able workers are less likely to draw a negative surplus. We also assume an additional and exogenous layoff shock with probability δ.

A displacement occurs with probability λ_2 times the probability of drawing $0 \leq z' < z$. This type of forced mobility is standard in the empirical search literature as, again, many moves are observed with lower wages. In the case of wage posting, the wage paid to the worker is fixed, independent of the hiring mode. In a sequential auction framework, how wages are determined in the case of a displacement is a question that has not received any attention. We assume that the hiring wage is equal to job productivity if displacement. Thus, there will not be any further wage increase on the job beyond the experience effect $\gamma(t)$.

Worker stays anyway. This is the mirror case of the preceding one. Now, the worker eventually decides to stay whatever the poacher bids. If $z' \leq z$, the incumbent wins and this is like in the normal poaching game. If $z' > z$, the worker should move, but does not. Perhaps the worker incurs a mobility cost that the auctioning employers are unaware of. As a consolation prize, she receives a wage increase up to the maximum wage the incumbent employer is willing to pay. This is a simple way of incorporating mobility costs into the wage competition model.

Specifically, with probability λ_3 a worker employed with match surplus z draws an alternative job offer with surplus z' resulting in the following wage:

$$w' = \max\left\{\gamma(t) + x + \min\left\{\beta z + (1-\beta)z', z\right\}, w\right\}.$$

If $z' > z$, the worker stays and receives the maximal wage ("best advancement") that the incumbent employer can pay.

To conclude, a displacement is a move that should not happen, a best advancement is a move that should happen but does not. We can think of Poisson rates λ_2 and λ_3 as a way of adding imperfections to the main auction theory, which allow for a decoupling between wages and job preferences without relying on measurement error, mobility costs or non financial job utility (amenities).

3 The steady state economy

We assume that the economy is stationary. By this we mean that all worker distributions remain fixed over time. This is useful for estimation purposes as it allows to use maximum likelihood to estimate the model with unobserved heterogeneity.

3.1 Market segmentation

We assume that the sampling distribution of firm types y has CDF $Q(y|x)$ and PDF $q(y|x)$. By conditioning the firm sampling distribution on the worker type, we allow for segmented search. Next, match types v are drawn independently of firm types from a distribution with CDF $H_i(v)$ and PDF $h_i(v)$, where $i = 0$ refers to non-employment and $i = 1, 2, 3$ refers to the various poaching modes.

Let $\overline{G}_i(x, z) = 1 - G_i(x, z)$ denote the probability that a worker x employed at a job with surplus z, draws a strictly more competitive alternative job offer. In particular, $\overline{G}_0(x, 0)$ is equal to the probability for a non-employee of drawing an acceptable job, i.e. such that match surplus z is positive. We have

$$\overline{G}_i(x, z) = \int \overline{H}_i(z - f(x, y))\, q(y|x)\, \mathrm{d}y,$$

where $\overline{H}_i(v) = 1 - H_i(v)$. Hence, $\overline{G}_i(x, z)$ is a non-increasing function of the current match surplus z.

Then we denote

$$g_i(x, z) = -\frac{\partial \overline{G}_i(x, z)}{\partial z} = \int h_i(z - f(x, y))\, q(y|x)\, \mathrm{d}y.$$

This is the instantaneous probability density, given x, of drawing a job (y, v) such that $f(x, y) + v = z$.

3.2 The wage and employment mobility process

Under these assumptions, a worker's state is entirely described by the following collection of variables: experience t, ability x, employment status $s \in \{0, 1\}$ (employed or not), and if employed ($s = 1$), the employer type y, the match surplus z, and the worker surplus $\zeta \in [0, z]$ mapping the wage one-to-one given t, x, z:

$$w = \gamma(t) + x + \beta z + (1 - \beta)\zeta. \tag{2}$$

The worker's state at any point in time (t, x, s, y, z, w) is a Markovian process. To see that, let (t, x, s', y', z', w') be a transition from $(t - 1, x, s, y, z, w)$. If $s' = 1$, we observe the wage $w' = \gamma(t) + x + \beta z' + (1 - \beta)\zeta'$, again being observationally equivalent to the worker surplus ζ'.

Normal firm-to-firm transition. The worker makes a firm-to-firm transition to a new firm y' ($s' = s = 1$ and $y' \neq y$), and her new wage is greater than the incumbent match's productivity: $w' > \gamma(t) + x + z$. Thus,

$$z' = \frac{1}{\beta}\left[w' - \gamma(t) - x - (1 - \beta)z\right] > z \quad \text{and} \quad \zeta' = z.$$

This occurs with probability density

$$\left[\lambda_1 q_1(y'|x) h_1(z' - f(x, y')) + \lambda_2 q_2(y'|x) h_2(z' - f(x, y'))\right] \mathrm{d}z'.$$

Whether the worker quits anyway does not matter as the alternative job is more productive ($z' > z$).

Displacement. The worker makes a firm-to-firm transition to a new firm y', and her wage is lower than the current productivity: $w' < \gamma(t) + x + z$. Thus,

$$z' = w' - \gamma(t) - x < z \quad \text{and} \quad \zeta' = z'.$$

This occurs with probability density $[\lambda_2 q_2(y'|x) h_2(z' - f(x, y'))] \, \mathrm{d}z'$.

Within-job wage increase. The worker stays with the same employer ($s = s' = 1$ and $y = y'$) and her wage increases to $w' \in (w, \gamma(t) + x + z]$. Then,

$$z' = z \quad \text{and} \quad \zeta' = \frac{1}{1-\beta} \left[w' - \gamma(t) - x - \beta z\right] \in (\zeta, z).$$

This occurs with continuous probability density $[\lambda_1 g_1(x, \zeta') + \lambda_3 g_3(x, \zeta')] \, \mathrm{d}\zeta'$ (given x, y, z). Whether the worker stays anyway does not matter as the alternative job is less productive.

Best advancement. The worker stays with the same employer ($s = s' = 1$ and $y = y'$) and her wage increases to the maximal wage $w' = \gamma(t) + x + z$. Thus,

$$z' = \zeta' = z.$$

This occurs with probability mass (given x, y, z) $\lambda_3 \overline{G}_3(x, z)$.

Layoff. The worker becomes not employed ($s = 1$ and $s' = 0$) with probability $\delta + \lambda_2 G_2(x, z)$.

Re-employment. The worker leaves non-employment ($s = 0$ and $s' = 1$) to a firm y' and her new wage is greater than her reservation wage: $w' \geq \gamma(t) + x$. Hence,

$$z' = \frac{1}{\beta} \left[w' - \gamma(t) - x\right] \geq 0 \quad \text{and} \quad \zeta' = 0.$$

This occurs with probability density is $\lambda_0 q_0(y'|x) h_0(z' - f(x, y')) \, \mathrm{d}z'$.

Staying not employed. The probability of remaining not employed one additional period ($s = s' = 0$) is $1 - \lambda_0 \overline{G}_0(x, 0)$.

Figure 1 gives an example of the type of wage trajectories that our model produces. The first spell (in blue) out of non-employment branches out into two possible second spells (in red), one with a wage increase and one with a wage cut.

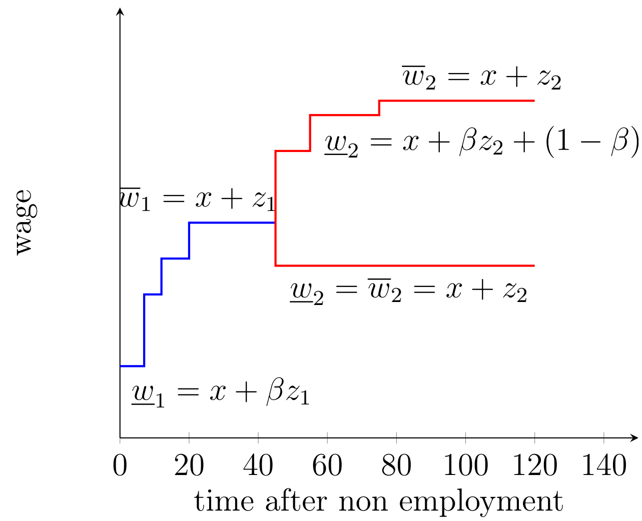


Figure 1: A worker trajectory with two employment spells following a non employment interruption (wage starts at minimum and ends at maximum; employer changes at $t = 45$, possibly upward ($z_2 > z_1$) or downward ($z_2 < z_1$))

The literature on wage dynamics has moved towards models characterized by increasingly richer and more flexible dynamics. These dynamics are usually interpreted as reflecting a sequence and accumulation of productivity shocks. However, the mechanisms by which productivity shocks are transmitted to wages is unclear. Thus, as Mincer equations incorporate more statistically sophisticated stochastic processes, their internal workings become increasingly opaque.

We can imagine ways of incorporating productivity shocks into our framework. For example, we could index wage contracts on productivity (as in the piece-rate models of Barlevy, 2008, Bagger, Fontaine, Postel-Vinay, and Robin, 2014). This implies a degree of employer commitment that may be difficult to justify. Or we could assume a total lack of commitment, and workers would need a credible alternative offer to benefit from productivity gains (see Lise, Meghir, and Robin, 2016). But adding productivity dynamics, for example on match-specific productivity v, would complicate the model by an order of magnitude and pose a further challenge for identification.

On the other hand, dynamic wage models usually neglect the firm heterogeneity dimension. The comparison with to two-sided unobserved heterogeneity now turns in our favor. Our model exhibits wage dynamics that are much more sophisticated than most two-way fixed-effects wage models (e.g. Abowd, Kramarz, and Margolis, 1999, Card, Heining, and Kline, 2013).

3.3 Steady-state distributions

The following distributions can be calculated assuming stationarity:

- the distribution of worker types, $\ell(t, x)$,
- the distribution of non-employees (including self-employed and unemployed) by worker type, $u(t, x)$,
- the distribution of match productivity components, $m(t, x, z)$,
- the distribution of wage components, $n(t, x, y, z, \zeta)$.

These distributions can be calculated by solving simple first-order differential equations. For example, Let $\ell_0(x)$ denote the distribution of new-born workers and let ξ be the retirement rate. The distribution of individuals by age and type is $\ell(t, x)$ such that

$$\frac{\partial \ell(t, x)}{\partial t} = -\xi \, \ell(t, x).$$

Hence,

$$\ell(t, x) = \ell_0(x) e^{-\xi t}. \tag{3}$$

The law of motion of $u(t, x)$ is

$$\begin{aligned}\frac{\partial u(t, x)}{\partial t} &= [\delta + \lambda_2 G_2(x, 0)] \, [\ell(t, x) - u(t, x)] - \left[\xi + \lambda_0 \overline{G}_0(x, 0)\right] u(t, x) \\ &= [\delta + \lambda_2 G_2(x, 0)] \, \ell(t, x) - d_0(x) u(t, x),\end{aligned}$$

with

$$d_0(x) = \xi + \delta + \lambda_0 \overline{G}_0(x, 0) + \lambda_2 G_2(x, 0).$$

Jobs terminate either with straight probability δ because of layoff, or because of a displacement shock with a negative surplus draw. Workers leave non-employment either because they retire or find a job. The solution to this first-order ODE with initial condition $u(0, x) = \ell_0(x)$ (i.e. assuming all workers start their career as a non-employee) is the cumulative sum of inflows from out of employment in time interval $[0, t]$ who have not yet exited non-employment by t. Specifically,

$$u(t, x) = \ell_0(x) e^{-d_0(x) t} + [\delta + \lambda_2 G_2(x, 0)] \int_0^t \ell(t', x) \, e^{-d_0(x)(t - t')} \, \mathrm{d}t'.$$

We can similarly calculate $M(t,x,z) = \int_0^z m(t,x,z')\,\mathrm{d}z'$ from the ODE (with 0 as initial condition):

$$\begin{aligned}\frac{\partial M(t,x,z)}{\partial t} &= \lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] u(t,x) \\ &\quad + \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right] [\ell(t,x) - u(t,x) - M(t,x,z)] \\ &\qquad - \left[\xi + \delta + \lambda_2 G_2(x,0) + \lambda_1 \overline{G}_1(x,z) + \lambda_2 \overline{G}_2(x,z)\right] M(t,x,z).\end{aligned}$$

The inflow is made of workers who were not in the stock $M(t,x,z)$ and enter; that is, the non-employee and the displaced workers with match surplus greater than z, who draw a new match surplus in $[0,z]$. The outflow comprises retirees (ξ), laid-off ($\delta+\lambda_2 G_2(x,0)$) and workers who are poached or displaced in jobs with a surplus greater than z ($\lambda_1 \overline{G}_1(x,z) + \lambda_2 \overline{G}_2(x,z)$).

We refer to Appendix B for the complete state variable distribution.

3.4 Identification

Proving identification from individual worker trajectories is complicated because the different wage components (ability x, match surplus z and worker surplus ζ) are not independent of each other and over time. However, we propose in Appendix A a heuristic argument that highlights what in the model is well identified and what is likely to rely on parametric restrictions. We now summarize this discussion.

We neglect the experience effect $\gamma(t)$. Let us simply assume that we can filter it out. Then, suppose that for each worker there exists a sequence of consecutive employment spells such that: 1) wages increase in all spells (i.e., for all spells, the last wage is strictly greater than the first wage), 2) the first spell follows a period of non-employment, 3) the last wage of the sequence is a best advancement wage.[7] We first show that, under these conditions, the sequence of minimum and maximum wages in each spell forms a system of equations that point identifies β, worker ability x and the sequence of match surpluses z and worker surpluses ζ after a non-employment spell.

Key for identification is that the first wage in each spell is only a function of worker ability x and of the match surpluses z and z' before and after an employment transition. A transition from non-employment further reduces the number of latent variables because the match surplus in non-employment is by definition zero. It also helps that the model allows the last wage in a spell to be equal to match productivity $x + z$. Suppose that firm competition is strong enough that wages increase to match productivity in each spell with

[7] In practice, we would infer that the last wage of an employment is equal to the match output from the relatively long duration of the last step.

a high probability, then the bargaining power coefficient β is identified from firm-to-firm transition wages.

Second, assuming that the worker state variables are known (i.e. worker ability x, firm technology y, match surplus z), the transition rates $\delta, \lambda_1, \lambda_2, \lambda_3$ are identified by transition probabilities (rate time the probability of moving) if $f(x, y)$ is bounded when y varies and if the supports of match components v are also bounded (with certain conditions on those bounds). Hence, match surpluses z are bounded from below and from above, and there exist matches that cannot be improved or that can only be improved. These rates are thus identified as maximum or minimum transition frequencies given the current state.

Third, identifying $f(x, y)$ from v in $z = f(x, y) + v$, knowing x, y, is a standard non-parametric censored regression problem, which is identified assuming for example that the distributions H_i have zero median.

The conclusion of this discussion is that we can understand how the data will impose important restrictions on the parameters. However, it seems difficult to estimate the model nonparametrically. In practice, we will impose parametric restrictions and use Maximum Likelihood for estimation.

4 Data

4.1 Firm characteristics

We use the DADS-POSTES dataset, covering the years 2015-2019, which is available via the secure CASD data center and contains all the annual salary declarations for all employees of all legal companies.[8] All firms are identified by a unique identifier, allowing to follow the evolution of their workforce over time. We keep all private, non-agricultural firms. A record line is any employee in any year. We only keep full-time employees aged 25-55. For each firm, we calculate its size as the number of records divided by 5. For each remaining employee, we calculate a daily wage as the recorded annual earnings divided by the number of days worked in the firm in the year.

We classify all worker*year observations into 4 skill groups: managers and engineers (corresponding to the 2-digit PCS categories 37 and 38,[9] technicians, supervisors, and other skilled trade and administrative workers (PCS 46, 47 and 48), unskilled care, trade and administrative workers (PCS 54, 55 and 56), and non-farming production workers (PCS 62 to 68).

[8]DADS stands for Déclarations Annuelles de Données Sociales.

[9]PCS (Professions and catégories socio-professionnelles) is a usual nomenclature of occupations.

Then, for each firm and by skill group, we calculate the mean, median, p25 and p75 of the distribution of all recorded daily wage over the period 2015-2019. We summarize these firm-level wage distributions further by the first two principal components. As expected, the first factor measures the wage location and the second factor measures the wage dispersion. We also use as a third measure the log of the average firm size over the period. It is provided in the data as a headcount of the number of employees at the end of each calendar year. We explain in section 5.1 how exactly these three measures are used.

4.2 Worker trajectories

DADS-POSTES is not a worker panel as it lacks worker identifiers. However, we can merge it — by firm ID and keeping the three firm characteristics calculated in the previous step (size, wage location and dispersion) — with the DADS-PANEL, which records wage and employment trajectories for all individuals born in October.

We divide the 5 years 2015-2019 into 10 semesters in order to get shorter time periods. This is more consistent with our continuous-time model and limits the somewhat arbitrary assignment rules one has to make for workers who move or change status several times in a given discrete period.

If a worker has multiple overlapping employment records in a given semester, we select the one with the highest number of paid working days. If the total number of paid days is less than 30, or if there is no employment record, we classify the worker as unemployed. We then keep prime-age workers (25-45 y.o. in 2015) that are always salaried full-time and always working in a private firm with at least 3 employees during the observation period. We further remove workers whose recorded occupation is farmer, CEO, craftsman or merchant at some point in the observation period. We finally assign each worker to the skill group (same definition as before) in which he or she is observed to be employed for the greatest number of periods. The wage definition is also the total net earnings received in the year divided by the total number of paid days. We remove individuals with extreme wages, extreme wage changes or extreme firm types at some point of their career trajectory. Additional details on these sample selection and cleaning steps are provided in Appendix E.

For each skill group, our estimation sample is a random draw of 50,000 men employed at least once over the five-year sample period. The following variables are recorded: worker ID, year, age, employed-unemployed, skill group, wage.[10]

[10] Note that a similar sample of women results in very high non-employment rates. It is not uncommon for more than 25% of female employees to be observed not employed in one given year (according to the above definition). We therefore decided to focus on men.

4.3 Noise reduction

Here, statistical noise takes the form of wage fluctuations which cannot be explained by the model: wages do not decrease within an employment spell, and they increase infrequently due to poaching as persistent discrete jumps. Unfortunately, we do see wages decrease, and we do not know exactly when wages increase because of promotions. One possibility would be to add a measurement error in the wage equations. However, because of unobserved heterogeneity and unobserved initial states, adding measurement error would seriously complicate the calculation of the likelihood. So instead of modeling noise, we preferred to filter it out ex ante in a transparent manner. It is common practice in statistics to apply noise filtering algorithms to the data, as it is in macro-econometrics (e.g. Hodrick-Prescott filter) and in micro-econometrics (e.g. transitory-permanent deconvolution).

We first estimate the experience effect $\gamma(t)$ by regressing log wages on age and age squared. This is not optimal, but filtering out ex ante all covariate effects considerably simplifies the estimation procedure. Then, employment spell by employment spell, we smooth individual wage trajectories using the following clustering procedure, which resembles single-linkage clustering. Only adjacent linkages are considered, and wages are updated iteratively with a stopping rule. The procedure is described as follows. Let the log wages within a job be denoted by $w_k, k = 1, ..., n$. We need to cluster the n periods into adjacent clusters within which the wage remains constant, and jumps upward between clusters. Initialize the cluster index as $c_k = k$. Let $\mathrm{d}w_k = w_k - w_{k-1}, k > 1$. We iterate the following algorithm:

1. Compute the smallest non-zero wage difference $\underline{\mathrm{d}w} = \min \{ \mathrm{d}w_i : \forall i,\, \mathrm{d}w_i \neq 0 \}$. Proceed to Step 2 if $\underline{\mathrm{d}w} < .05$ (wage drop or small wage increase). Note that this 5% threshold is arbitrary. Otherwise, stop.

2. Let k^* be the smallest index such that $\underline{|\,\mathrm{d}w|} = \min \{ |\,\mathrm{d}w|_i : \forall i, dw_i \neq 0 \}$ (smallest non-zero absolute wage difference). Update the cluster index as follows. For all k such that $c_k = c_{k^*}$, update c_k to c_{k^*-1}. That is, combine the two different clusters that contain indices k^* and $k^* - 1$, respectively.

3. Update the wage of the combined cluster as the average wage (sum of wages in both clusters divided by the corresponding number of observations). Go to #1.

This procedure combines clusters until all wage steps are positive and exceed the threshold, or all wage steps are zero. The procedure is mean preserving. A numerical example of the algorithm is given in Table 1.

Now, after running the noise reduction algorithm, suppose that we observe two consecutive employment spells with a wage cut at the transition. This can only happen in case of a

Table 1: The clustering and wage filtering algorithm on an example with a wage threshold equal to 3

	initial			1			2			3			4		
k	w	dw	c	w	dw	c	w	dw	c	w	dw	c	w	dw	c
1	5		1	5		1	3.9		1	4.5		1	4.5		1
2	2.8	-2.2	2	2.8	-2.2	2	3.9	0	1	4.5	0	1	4.5	0	1
3	5.8	3	3	5.8	3	3	5.8	1.9	3	4.5	0	1	4.5	0	1
4	9.8	4	4	10.3	4.5	4	10.3	4.5	4	10.3	5.8	4	11.2	6.7	4
5	10.8	1	5	10.3	0	4	10.3	0	4	10.3	0	4	11.2	0	4
6	13	2.2	6	13	2.7	6	13	2.7	6	13.0	2.7	6	11.2	0	4

displacement. But then, the wage should remain constant after a displacement. Therefore, we operate another round of regularization. Suppose that a spell suspected to be a displacement nevertheless shows n increasing steps. We start by replacing the first two steps by the mean. If the wage transition remains negative, we continue with the aggregation of the first three steps, and so on until we have exhausted all the steps.

The raw log wages have a variance of 0.217 (4.943mi observations). After keeping only men, the variance is 0.217 (3.370mi obs). After winsorizing wages, firm types and wage variations, the variance is down to 0.164 (3.004mi obs). Then, keeping only prime-aged workers aged 25-45, the variance is 0.147 (2.051mi obs). Finally, using a 5% threshold, the regularized log wage has a variance of 0.144 and the noise variance is therefore $0.147-0.144 = 0.003$ (or 1.90% of the total variance). The correlation between the raw and regularized log wage is $\sqrt{1-0.0190} = 0.990$. The noise reduction is a lot smaller than the residual variance in two-way fixed effects models' estimations (Abowd, Kramarz, and Margolis, 1999, and followers). Appendix E details the effects on the distribution of wages of these various data selections.

5 Estimation method

5.1 Estimation method

Our estimation procedure is similar in many ways to that of Bonhomme, Lamadon, and Manresa (2019) (BLM). They proposed a two-stage estimation procedure. First, they use k-means to classify firms from firm-level data on wages and size. This relies on the property that, in equilibrium, firms of the same type should display similar wage distributions (across their employees). Second, they take the classification of firms as given and maximize the likelihood of worker wage and employment trajectories. Under the finite mixture assumption, this can be easily done using the Expectation-Maximization (EM) algorithm. Lentz, Piyapromdee, and Robin (2023) (LPR) propose a slightly different approach. Instead of using some firm-level statistics to pre-classify firms, they use the likelihood of worker data to classify the firms into a finite number of types. They proceed iteratively, alternating the firm classification and the EM stage.

Our model assumes continuous worker and firm heterogeneities. We cannot apply the previous approaches directly, but we can draw some inspiration from them. We postulate that, in equilibrium, the firm type y is well approximated by a linear combination of observed firm characteristics. Specifically, $y = \sum_{\ell=1}^{3} \alpha_\ell y_\ell$, where y_1, y_2, y_3 are the three firm characteristics described in Subsection 4.1 (wage location, wage dispersion and size). As BLM, we postulate that certain firm characteristics are good predictors of the observed firm type. But, as LPR, we estimate the meta-parameters α_ℓ by Maximum Likelihood together with the other structural parameters. For identification, we normalize $\alpha_1 = 1$ (the wage location parameter). We found in simulations (see Appendix D) that the equilibrium link between wage location and the type y is not linear. Adding wage dispersion, the link between the predicted type and the true type is not exactly on the 45% line. Adding firm size, the prediction is almost perfect.[11]

5.2 Parametric specification for the structural model

The match surplus $f(x, y)$ is defined as

$$f(x,y) = c\left[y - \frac{1}{2b}(y - ax)^2\right],\ y \geq 0,$$

for some $a, b, c > 0$. This specification implies that log match production, $x + f(x, y)$, has an additive worker effect x, an additive firm effect cy and a cost $\frac{c}{2b}(y - ax)^2$. Gautier and Teulings (2015) develop an equilibrium search model (with on the job search) assuming a similar specification of match productivity (see also Boerma, Tsyvinski, Wang, and Zhang, 2023).

One can easily verify that $f(x, y)$ is supermodular:

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{ac}{b} \geq 0.$$

Moreover complementarity decreases with b. At the lowest limit ($a = 0$ or $b = +\infty$) $f(x, y)$ does not depend on x and $f(x, y)$ is like a firm effect.

[11]Our estimation method is also linked to the Krusell-Smith model (Krusell and Smith (1998)).

The optimal firm type for a worker x is

$$y^*(x) = \arg\max_y f(x,y) = ax + b.$$

The optimal surplus is $f(x, y^*(x)) = c\left(ax + \frac{b}{2}\right)$. Hence, a worker of type x produces $\exp\left[\gamma(t) + x\right]$ by himself and a maximum of $\exp\left[\gamma(t) + (1+ac)x + \frac{bc}{2}\right]$ when matched optimally.[12] It follows that

$$\begin{aligned} f(x,y) &= c\left[y - \frac{1}{2b}(y-ax)^2\right] \\ &= c\left(ax + \frac{b}{2}\right) - \frac{c}{2b}(y - ax - b)^2 \\ &= cy - \frac{c}{2b}y^2 + \frac{ca}{b}xy - \frac{ca^2}{2b}x^2 \end{aligned}$$

(optimal surplus minus distance to optimal match).

The parameters to estimate are

- transitions rates ξ (retirement rate), δ (exogenous layoff rate) and job offer arrival rates $\lambda_0, \lambda_1, \lambda_2, \lambda_3$;
- the experience effect $\gamma(t)$ modeled as a stepwise constant function of age;
- the distribution of worker heterogeneity $\ell_0(x)/L_0 = \mathcal{N}(\mu, s^2)$ (normal);
- the parameters of the sampling distribution of firm types $q(y|x) = \mathcal{N}(\kappa + \rho(x-\mu), \omega^2)$, common to all job offers;
- the surplus function parameters a, b, c;
- the bargaining power coefficient β;
- the distribution of job-specific effects $h_i(v) = \mathcal{N}(0, \sigma_i^2)$, $i = 0, 1, 2, 3$. To help identification, we force $\sigma_3 = \sigma_1$.

The sampling distribution of match surpluses is therefore

$$g_i(x,z) = \frac{1}{2\pi}\int \frac{1}{\sigma_i}\frac{1}{\omega}\exp\left(-\frac{1}{2\sigma_i^2}\left[z - f(x,y)\right]^2 - \frac{1}{2\omega^2}\left[y - \kappa - \rho(x-\mu)\right]^2\right)\,\mathrm{d}y.$$

[12] In general, if f is assumed concave with respect to y ($f_2'' < 0$), for any x there exists an optimal match $y^*(x) = \arg\max_y f(x,y)$. In addition, if f is supermodular ($f_{12}'' > 0$), then $y^*(x)$ is increasing.

We shall further constrain $\sigma_1 = \sigma_3$. We also pre-estimate ξ as the inverse of the average experience of workers in the sample (in weeks or months), as the PDF of experience is exponential $\xi e^{-\xi t}$.

5.3 Maximum Likelihood

An individual observation is a sequence $(t_n, s_n, d_n, y_n, w_n, n = 1, ..., N)$ where $t_n = t_1 - 1 + n$ is age or experience, $s_n = U, E, R$ is the employment status (non-employee, employed or retired), $d_n \in \{0, 1\}$ is the indicator of job mobility, y_n is the employer's type and w_n is the wage in period n (both missing if not employed). Initially, the worker can be employed or not. If employed, her first log-wage is of the form

$$w_1 = \gamma(t_1) + x + \beta z_1 + (1 - \beta)\zeta_1,$$

where $z_1 = f(x, y_1) + v_1$. Note that neither x, nor v_1, nor ζ_1 are observed. However, under the stationarity assumption, we can calculate the joint distribution of all the wage components: $n(t_1, x, y_1, z_1, \zeta_1)$. The calculation of the steady-state distributions is tedious, but it pays in the end as it allows to use maximum likelihood to estimate the model. We will have to integrate out two out of the three unobserved variables x, v_1, ζ_1. In practice, we substitute out z_1 given w_1 and integrate the likelihood over x and ζ_1. Appendix C details the likelihood.

6 Results

6.1 Parameter estimates

Parameter estimates are displayed in Table 2. We note that δ is estimated to be negligible. This is because displacements with a negative surplus draw are sufficient to predict all layoffs. Wages change according to the various Poisson processes of alternative job draws. We estimate λ_3 (best advancements) much larger than λ_1 (normal auctions) and λ_2 (displacements), and slightly less than λ_0 (offers to non-employee workers).

The firm heterogeneity index depends positively on wage dispersion for managers, but negatively for administrative and production workers. Size has a negative effect, except for production workers.

We also note that c, which sets the level of $f(x, y)$ in surplus z, seems small. The worker surplus sharing rule β is also estimated rather small, near .15. Such a low bargaining power is not surprising.[13] We can hardly compare our estimates of bargaining power with those

[13]Cahuc, Postel-Vinay, and Robin (2006) already found, using same French data, very small estimates,

Table 2: Parameter estimates

		Managers		Supervisors		Admin		Production	
		Estimate	Std	Estimate	Std	Estimate	Std	Estimate	Std
rates	δ	.004	.001	.006	.001	.020	.001	.003	.001
	λ_0	.364	.005	.303	.007	.272	.006	.420	.004
	λ_1	.010	.002	.054	.005	.111	.013	.029	.003
	λ_2	.058	.001	.021	.001	.030	.001	.061	.001
	λ_3	.284	.008	.260	.014	.190	.017	.357	.012
directed search $q(y\|x)$	κ	.587	.010	.521	.014	.160	.017	.633	.011
	ρ	2.137	.014	3.587	.025	4.194	.034	3.789	.021
	ω	.786	.003	.769	.003	.875	.005	1.138	.003
firm type y	α_2	.613	.011	.000	.011	-.160	.012	-.320	.009
	α_3	-.008	.001	-.041	.002	-.030	.003	.010	.002
match production $f(x,y)$	a	.486	.098	1.003	5.516	.618	.237	1.009	.142
	b	3.577	1.229	30.742	371.98	1.693	1.155	2.844	.941
	c	.047	.003	.007	.004	.015	.006	.028	.003
bargaining coeff.	β	.168	.001	.145	.002	.162	.002	.160	.001
worker heterogeneity $\ell_0(x)$	μ	3.923	.002	3.673	.002	3.426	.003	3.468	.002
	s	.267	.001	.208	.001	.215	.001	.222	.001
match shocks $h_i(v)$	σ_0	1.487	.015	1.245	.020	1.275	.022	1.344	.013
	σ_1	.216	.002	.179	.003	.195	.004	.212	.002
	σ_2	.478	.003	.446	.007	.394	.006	.538	.004
	σ_3	.216	.002	.179	.003	.195	.004	.212	.002

in the rest of the literature because the wage ladder mechanism is precisely a rent-sharing mechanism. Workers can in principle acquire a lot of the match surplus through Bertrand competition only (when $\beta = 0$).

Finally, we find that σ_0 is a lot bigger that $\sigma_1 = \sigma_3$ and σ_2. This points at a large reseting effect of non-employment.

6.2 Model fit

Table 3 displays sample statistics along with their predictions by model simulation. Mobility, whether in employment or in wages, is a rare event. Wage mobility is more frequent within than between employment spells. The model does a good job at fitting the state and mobility data, although it tends to predict too big wage changes across consecutive employment spells. This is because λ_1 adds to λ_2 to generate moves with a wage increase.

Figure 2 shows how the model fits employer type and wage distributions for the whole population (the four skill groups pulled together). Fits of distributions for skill groups considered separately share similar features.[14] The distribution of employer types (across workers) is well fitted. The overall wage distribution displays a longer right tail in the simulation. The distribution of wage changes within an employment spell is well fitted. The distribution of wage changes across two consecutive employment spells shows an undesirable peakedness (a spike and longer tails) in the simulation.

We then illustrate the respective roles of the four different wage-setting cases. Figure 3 shows the simulated distributions of wage increases for normal wage increases ($\lambda_1 + \lambda_3$, $z' < z$), best advancements (λ_3, $z' > z$), normal hiring ($\lambda_1 + \lambda_2$, $z' > z$) and displacements (λ_2, $z' < z$).[15] Best advancements increase the skewness because they come after a series of wage increases. Wage increments should decrease with the number of steps. Forced moves (displacements) generate a very dispersed distribution of wage changes over both negative and positive values. This is likely a consequence of the normal distribution. In order to generate the right frequency of displacements for average current match surplus values, a large dispersion of zero-mean match-specific components (σ_2) is needed.

We next examine the different frequencies of these four cases.

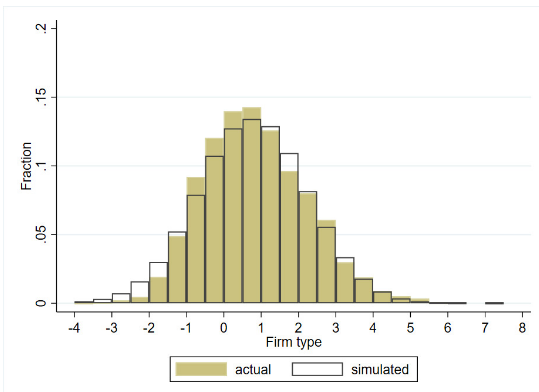
generally between 0 and .30, depending on the firm's sector and the employee's skill, and Bagger, Fontaine, Postel-Vinay, and Robin (2014) found .30 on Danish data.

[14]See Figures F.2 to F.5 in Appendix F.

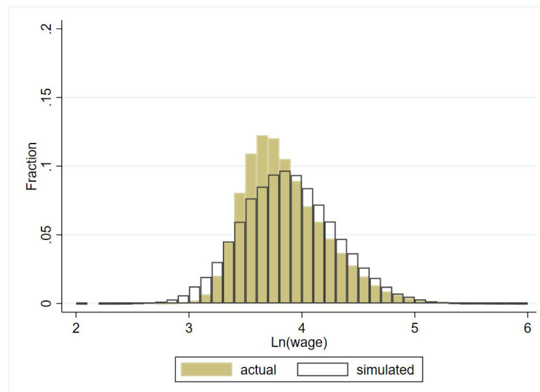
[15]See Figures F.6 to F.9 for distributions by skill group.

Table 3: Descriptive statistics, actual and fit

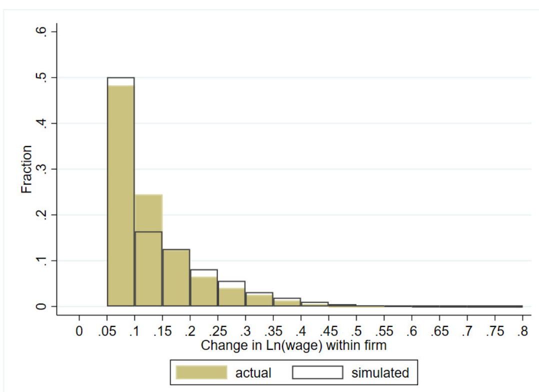
	Managers and engineers				Overseers and technicians				Admin personnel				Production workers			
	Actual		Fit		Actual		Fit		Actual		Fit		Actual		Fit	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Share employed	89.2%		84.4%		94.3%		90.1%		79.6%		79.0%		88.1%		85.3%	
Across two consecutive years																
Stay unemployed	9.2%		12.5%		4.8%		8.4%		17.6%		18.2%		9.8%		11.7%	
Leave non-employment	1.8%		2.8%		1.0%		1.5%		3.1%		2.8%		2.3%		3.1%	
Lose employment	1.2%		2.7%		0.6%		1.5%		2.2%		2.7%		1.6%		3.0%	
Stay employed	87.8%		82.0%		93.5%		88.6%		77.0%		76.3%		86.3%		82.3%	
Daily wage (in euros)	73.51	30.91	76.47	29.68	54.63	13.92	56.23	16.63	42.96	11.99	43.50	13.45	46.11	13.79	47.16	17.12
Log daily wage	4.23	0.33	4.27	0.36	3.97	0.23	3.99	0.27	3.73	0.25	3.73	0.29	3.79	0.27	3.80	0.32
Log wage after UE move	4.24	0.39	4.11	0.31	3.83	0.30	3.82	0.24	3.56	0.25	3.59	0.25	3.58	0.27	3.63	0.26
Log wage after JJ move	4.22	0.28	4.24	0.36	3.91	0.21	3.92	0.30	3.66	0.21	3.68	0.28	3.74	0.26	3.79	0.33
Firm type	0.67	0.91	0.65	0.96	0.56	1.03	0.53	1.07	0.32	1.20	0.22	1.24	0.90	1.32	0.71	1.40
Firm mean wage	0.92	0.93	.	.	0.80	1.03	.	.	0.45	1.19	.	.	0.79	1.30	.	.
Firm wage sd	-0.32	0.24	.	.	-0.18	0.27	.	.	-0.16	0.33	.	.	-0.11	0.31	.	.
Firm log size	6.42	2.30	.	.	6.00	2.27	.	.	5.50	2.92	.	.	5.34	2.54	.	.
Employed in two consecutive periods																
Same firm, same wage	88.0%		92.1%		91.0%		94.5%		91.3%		92.5%		88.8%		91.4%	
Same firm, wage increase	8.0%		4.8%		7.3%		4.2%		6.9%		5.3%		7.2%		5.3%	
Different firm, wage increase	2.0%		1.7%		0.8%		0.9%		0.8%		1.3%		1.8%		2.0%	
Different firm, wage drop	2.0%		1.4%		0.8%		0.5%		1.0%		0.8%		2.2%		1.3%	
Wage change																
All	0.9%	5.3%	0.5%	5.9%	0.7%	3.5%	0.4%	3.5%	0.7%	3.9%	0.5%	4.3%	0.7%	5.2%	0.6%	6.0%
Same firm, wage increase	13.3%	8.0%	12.1%	10.0%	10.8%	5.4%	9.2%	7.9%	11.5%	6.0%	10.3%	8.6%	12.6%	7.4%	11.3%	9.4%
Different firm, wage increase	10.7%	8.8%	14.9%	16.0%	9.0%	7.2%	10.6%	14.0%	10.9%	8.2%	10.5%	12.6%	12.5%	9.8%	16.2%	18.4%
Different firm, wage drop	-13.4%	10.5%	-25.6%	19.8%	-10.8%	8.4%	-21.9%	16.4%	-13.1%	9.9%	-21.7%	16.7%	-15.9%	11.7%	-25.0%	20.2%



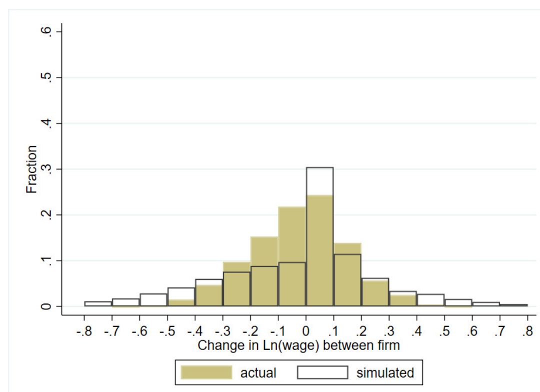
(a) All employer types



(b) All wages

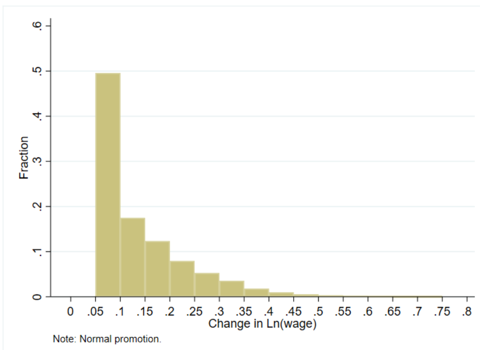


(c) Within-firm wage growth

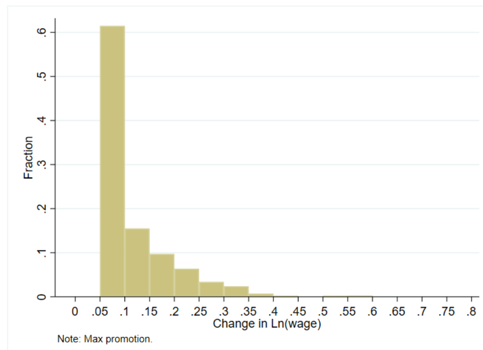


(d) Between-firm wage growth

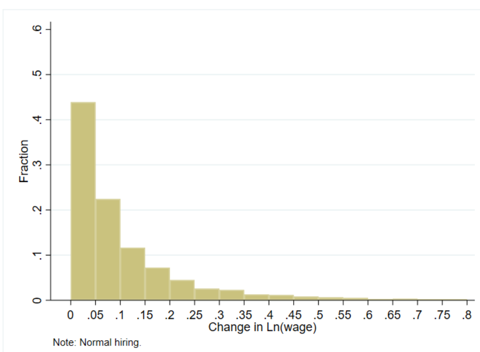
Figure 2: Fit of wage distributions, whole sample



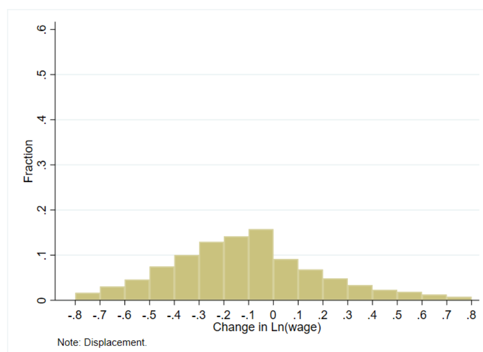
(a) Normal wage increases



(b) Best advancements



(c) Normal hiring



(d) Displacements

Figure 3: Distributions of simulated wage changes, whole sample

Table 4: Proportions and tenure by last event type

	Manager, engineers			Overseer, technician			Admin staff			Production worker		
	N	Tenure		N	Tenure		N	Tenure		N	Tenure	
	(%)	(mean)	(sd)	(%)	(mean)	(sd)	(%)	(mean)	(sd)	(%)	(mean)	(sd)
Last event unknown	4.06			15.19			5.74			4.07		
Non employed	15.13	5.34	4.67	9.90	6.27	5.40	20.64	7.05	5.94	14.27	4.70	4.02
Employed	80.81	10.73	7.66	74.91	12.94	8.23	73.62	11.26	7.90	81.66	10.50	7.63
Among employees												
Re-employment	16.02	4.92	4.28	10.74	5.18	4.56	20.07	5.32	4.56	15.55	4.46	4.04
Normal promotion	35.95	13.87	7.21	51.33	15.74	7.48	48.30	14.90	7.33	38.42	13.51	7.25
Best advancement	18.84	15.00	7.21	17.45	16.78	7.19	9.91	15.33	7.14	16.58	14.46	7.25
Normal hiring	10.13	8.30	6.76	8.04	10.64	7.74	9.52	8.59	7.05	11.31	8.41	6.92
Displacement	19.05	8.83	7.15	12.44	11.49	8.16	12.20	9.62	7.56	18.13	8.73	7.09

6.3 Inefficient mobility

The exogenous layoff rate δ is essentially zero. This is because the main source of job destruction happens to be the displacements with a negative surplus draw $z' < 0$. For the same reason, many surplus draws when unemployed will be negative. The reemployment rates are thus much lower than indicated by parameter λ_0.

The two additional sources of friction that we have introduced to the sequential auction mechanism (λ_2: the worker quits anyway, λ_3: the worker stays anyway) turn out to be predominant, and λ_1 is almost negligible. This is particularly true for interrupted poaching (λ_3): a large proportion of meetings with more competitive vacancies ($z' > z$) will not lead to a move, but instead to the wage being increased to its maximum value. This is because most job spells will start with a few early wage increases, and the final wage remains fixed for a long time until the next firm-to-firm move. Second, λ_2 is high because many firm-to-firm moves suffer from a wage cut. Finally, λ_1 is estimated to fit either the excess frequency of early wage increases within a job or the excess frequency of firm-to-firm moves with a wage increase (for higher skilled workers).

The general conclusion is that mobility is largely inefficient.

To evaluate the share of inefficient moves, we show in Table 4 the simulated proportions and tenures by last event type. We simulate trajectories using the estimated model for 30 periods (a period is equal to half a year). Then we take the last simulation periods and we classify workers by the last state change that happened to them: if unemployed, the last change was a layoff; if employed, the last change can be returning from non employment, or moving up the internal wage ladder (via external contacts of types λ_1 or λ_3 associated with a surplus draw $z' < z$), or moving to the top of the internal ladder (λ_3 contact with draw $z' > z$), or moving up the external ladder (contacts of λ_1 or λ_2 type with a surplus draw

Table 5: Type and wage sorting

	Managers	Supervisors	Admin	Production
corr(x, w)	.7780	.7623	.7812	.7388
corr(y, w)	.4795	.5383	.5697	.4646
corr(x, y)	.5901	.7022	.7194	.5992
No market segmentation ($\rho = 0$)				
corr(x, w)	.7621	.7517	.7569	.7051
corr(y, w)	.0255	.0017	.0080	.0379
corr(x, y)	-.0018	.0007	-.0018	.0060

Notes: The variables x, y, w used to calculate these correlations are all from a simulated economy (see Appendix D).

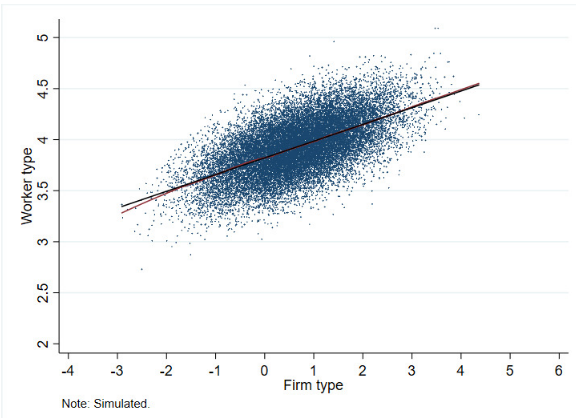
$z' > z$), or moving down the external ladder (λ_2 contact with $z' < z$).

A sizable proportion of workers are not employed. This is because non-employment includes unemployment, inactivity, self-employment, public employment and part-time work. Non-employment duration is estimated to be around 5-7 months with no clear ranking between the four different skill categories. As for employees, about 13% of them have returned from non-employment 4-5 months ago. A large proportion (35-50%) were most recently promoted internally. Having already been promoted makes another promotion less likely; this also implies that the current surplus z is relatively high. This selection effect explains the higher tenure. The proportion of workers paid their match productivity is around 10-19% (slightly less for admin workers), which is substantial. Their average tenure is higher. Finally, a transition to a job with a higher surplus is rather rare (about 8-11% of the time), while external mobility down the surplus ladder is much more common (about 12-19%). The average duration since the last job change is about 7-8 months.

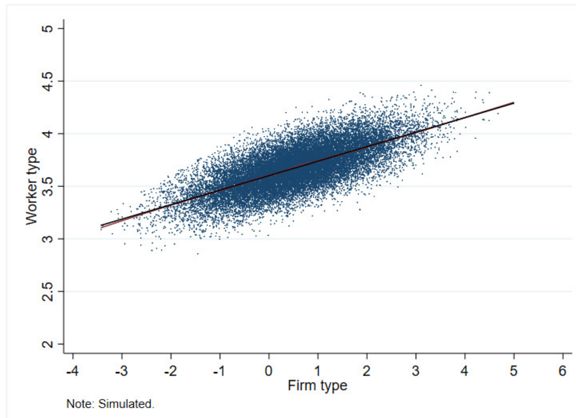
Overall, between job-to-job mobilities that should have occurred ($z' > z$) but did not, and those that did occur but should not have ($z' < z$), we estimate that the proportion of employees whose last state change is inefficient to be in the range 22-38%.

6.4 Type sorting and wage sorting

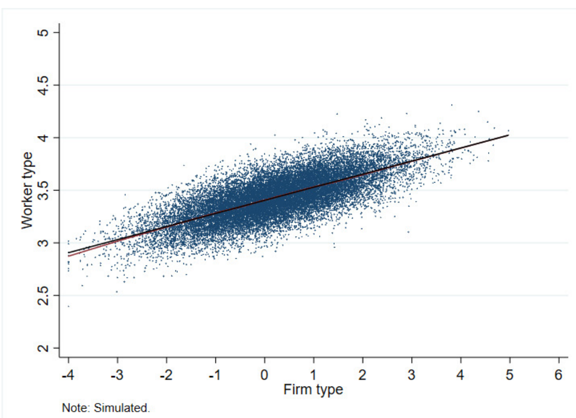
Workers draw offers according to different Poisson processes with different rates λ_i, but from the same distribution of firm types $q(y|x) = \mathcal{N}(\kappa + \rho(x - \mu), \omega^2)$. The parameter ρ is estimated very large, which indicates a strong directed search feature. Figure 4 shows the scatterplot of (x, y) observations. The association between workers' and firms' types is very strong, with correlations in the 60-70% range. Correlations between log wages and worker and firm types are also strong, stronger with worker types than with firm types though (see Table 5).



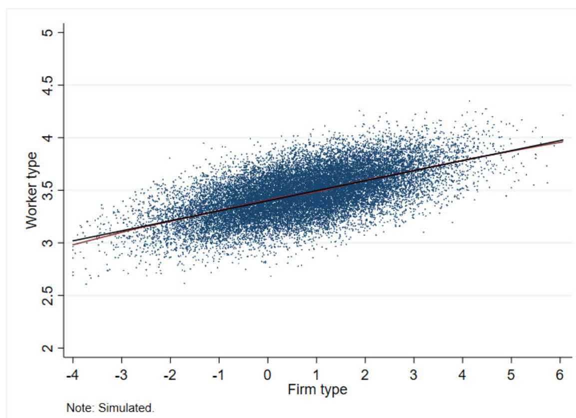
(a) Managers, engineers



(b) Overseers, technicians

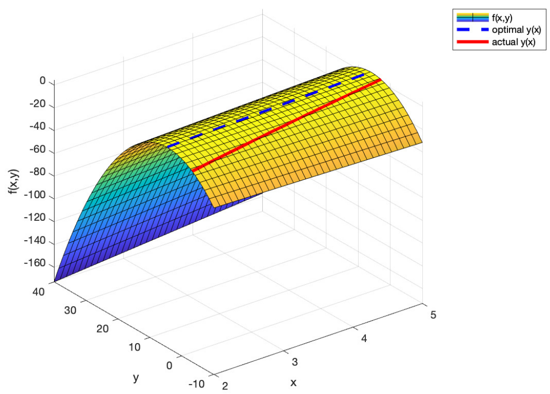


(c) Admin staff

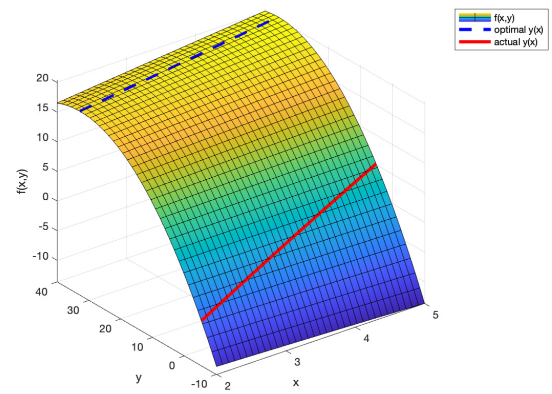


(d) Production workers

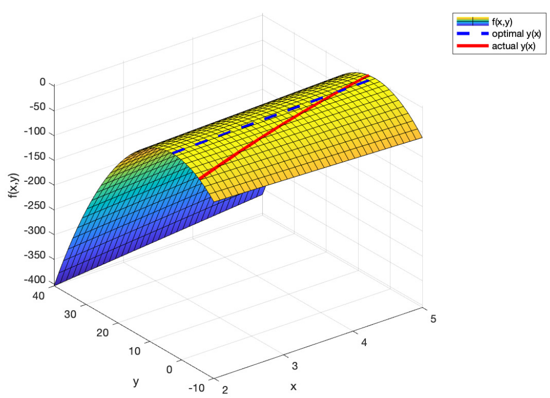
Figure 4: Cross-sectional distribution of worker and firm types



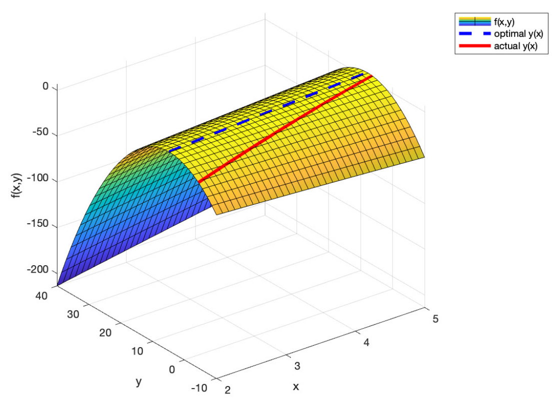
(a) Managers, engineers



(b) Overseers, technicians



(c) Admin staff



(d) Production workers

Figure 5: Match production function $f(x, y)$. The red line is the regression of y on x in the data. The dashed line is $y = ax + b$ and indicates the location of optimal matches

The match production $f(x, y)$ is proportional to $f(x, y) = y - \frac{1}{2b}(y - ax)^2$. The parameters a and b are imprecisely estimated, which means that complementarities between worker and firm types are not well identified. However, taking their estimated values seriously, we see in Figure 5 that $f(x, y)$ is increasing in y for all x and all skills. We plot on the $(x, y) \mapsto f(x, y)$ surface the location of the matching lines (i.e. the regression lines through the scatterplots in Figure 4). Apart from skill group 2 (supervisors), for which parameter b is very badly identified, realized matches are consistent with some production optimization.

Finally, we ran counterfactual simulations where we shut down all dependency to worker type in firm type draws. The equilibrium correlation between worker and firm types drops to 0 (see Table 5).

The conclusion is that there is strong worker-firm sorting, but it does not show up in wages. In the next section we investigate in details the role of the dynamic of match-specific wage shocks.

6.5 Match production function and match-specific surplus

Recall the wage equation

$$w = \gamma(t) + x + \beta z + (1 - \beta)\zeta,$$

where z is the current surplus and ζ differs according to the transition history: (i) $\zeta = 0$ if the last move is from non-employment; (ii) $\zeta \in (0, z)$ if the negotiation was normal (within-job or between-job wage increase); (iii) $\zeta = z$ in case of a best advancement or a displacement. Let us rewrite this equation as follows

$$w = \gamma(t) + x + m + u,$$

where $m = \beta z$ and $u = (1 - \beta)\zeta$ in the first two cases and $m = z$ and $u = 0$ in the last case. In this decomposition, u is the residual surplus contribution, which includes the part that is history dependent and presumably imperfectly correlated with the current match surplus.

Table 6 shows the variance decomposition of log wages into current match and residual contributions. The contribution of the worker effect is found to be close to 50%. The current match surplus explains about 25%, and the residual accounts for 20%. In addition, the covariance terms are negligible. These decompositions are not very different from the AKM-type wage variance decompositions.

Then, we turn to the decomposition of the variance of the match and worker surpluses (see Table 7). We find that the match output $f(x, y)$, including the firm effect, is negligible. Only the match-specific effect v seems to matter. Therefore, firm effects contribute particularly

Table 6: Total log-wage variance decomposition, by current surplus and past dependence

	Managers		Supervisors		Admin		Production	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
$\ln w$	4.258	.1284	3.978	.0792	3.720	.0858	3.790	.1057
x	3.929	.0704	3.676	.0441	3.433	.0459	3.468	.0495
		(54.8%)		(55.7%)		(53.5%)		(46.8%)
current match m	0.223	.0336	0.176	.0213	0.179	.0201	0.209	.0325
		(26.2%)		(26.9%)		(23.4%)		(30.7%)
residual u	0.107	.0225	0.126	.0181	0.108	.0154	0.112	.0230
		(17.5%)		(22.9%)		(17.9%)		(21.8%)
$2\text{cov}(x, m)$		.0028		.0001		.0017		.0033
$2\text{cov}(x, u)$		.0048		.0014		.0049		.0051
$2\text{cov}(m, u)$		-.0055		-.0058		-.0023		-.0077

Notes: The wage variance decomposition uses all wage observations (all four cases). The current match component is the contribution of the match surplus. The residual component is the residual history component.

Table 7: Match and worker surplus variance decomposition

	Managers		Supervisors		Admin		Production	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Match surplus z								
z	0.811	.6363	0.777	.4818	0.825	.5213	0.785	.5219
$f(x, y)$	0.015	.0037	0.002	.0001	-0.018	.0014	-0.026	.0055
v	0.796	.6393	0.775	.4819	0.842	.5229	0.811	.5272
Worker surplus ζ								
ζ	0.232	.0430	0.221	.0275	0.183	.0262	0.233	.0445
$f(x, y)$	0.017	.0030	0.002	.0001	-0.012	.0011	-0.015	.0043
v	0.215	.0425	0.219	.0274	0.195	.0274	0.248	.0476

Notes: The current surplus decomposition refers to all wage observations (all four cases). The current match component is the contribution of the match surplus. The residual component is the residual history component.

Table 8: AKM wage variance decomposition

	All		Managers		Supervisors		Admin		Production	
	Level	Share	Level	Share	Level	Share	Level	Share	Level	Share
					True data					
Var($\ln w$)	0.1430		0.1107		0.0563		0.0653		0.0736	
Var(worker)	0.1423	0.9948	0.1062	0.9588	0.0774	1.3745	0.0715	1.0943	0.0784	1.0656
Var(firm)	0.0183	0.1276	0.0131	0.1181	0.0279	0.4953	0.0196	0.2997	0.0197	0.2683
2Cov(wkr,firm)	-0.0220	-0.1537	-0.0133	-0.1197	-0.0517	-0.9189	-0.0293	-0.4489	-0.0290	-0.3946
Var(residual)	0.0041	0.0287	0.0040	0.0365	0.0023	0.0413	0.0034	0.0522	0.0041	0.0562
					Simulated data					
Var($\ln w$)	0.1811		0.1334		0.0787		0.0854		0.1096	
Var(worker)	0.1775	0.9802	0.1280	0.9592	0.0896	1.1380	0.0912	1.0676	0.1049	0.9564
Var(firm)	0.0169	0.0932	0.0151	0.1132	0.0178	0.2262	0.0127	0.1492	0.0141	0.1283
2Cov(wkr,firm)	-0.0249	-0.1375	-0.0185	-0.1386	-0.0328	-0.4161	-0.0236	-0.2768	-0.0178	-0.1627
Var(residual)	0.0116	0.0641	0.0088	0.0661	0.0041	0.0519	0.0051	0.0600	0.0086	0.0781

little to wage dispersion, but we find strong evidence of random match effects.

Lastly, we proceed to the estimation of an AKM two-way fixed-effects regression on our data and compare it with a similar estimation on simulated data (see Table 8. Both coincide, which confirms that the model fits the data well. However, the AKM variance decomposition delivers a much larger contribution of the worker effect and a sizable sorting effect (2 times the covariance between the worker and the firm effect) as in the original paper of Abowd, Kramarz, and Margolis (1999).[16] For an obscure reason, the match effect translates for one part into an augmented worker effect, a moderate firm effect and a negative sorting effect (2 times the covariance between the worker and the firm effect).[17]

6.6 Hiring wages

Finally, in Table 9 we show the same variance decomposition but for hiring wages, i.e. wages just after a firm-to-firm transition. The contribution of the current match has increased substantially, and the residual variance is now almost entirely compensated by the covariance between the current match component m and the historical residual u. Note that the covariance term is negative because in the case of maximum promotions and displacements, the current surplus term is higher and the residual is smallest. As in Di Addario, Kline, Saggio,

[16] Note that Babet, Godechot, and Palladino, 2022 estimate a smaller contribution of the worker effect and a small, positive sorting component. This is due to to the fact that they use hourly wages and all contracts, and we use daily wages and full-time contracts (see also Figure A1 in the Appendix B of their paper). By keeping full-time contracts, we reduce the dependence of wages to hours, which we do not model.

[17] The divergence between type sorting and wage sorting was already emphasized by Lentz, Piyapromdee, and Robin (2023).

Table 9: Log-wage variance decomposition, all hiring wages

	Managers		Supervisors		Admin		Production	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
$\ln w$	4.219	.1285	3.957	.1035	3.673	.0782	3.792	.1118
x	3.924	.0713	3.686	.0481	3.444	.0425	3.484	.0504
		(55.5%)		(46.5%)		(54.3%)		(45.1%)
current match m	0.223	.0477	0.202	.0429	0.151	.0332	0.220	.0489
		(37.1%)		(41.4%)		(42.5%)		(43.7%)
residual u	0.072	.0181	0.070	.0188	0.078	.0132	0.088	.0254
		(14.1%)		(18.2%)		(16.9%)		(22.7%)
$2\text{cov}(x, m)$		.0071		.0052		-.0011		-.0010
$2\text{cov}(x, u)$		.0004		.0044		.0032		.0060
$2\text{cov}(m, u)$		-.0163		-.0160		-.0128		-.0178

Table 10: Log-wage variance decomposition, hiring wages, normal negotiation ($\Delta z > 0$)

	Managers		Supervisors		Admin		Production	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
$\ln w$	4.205	.1110	3.921	.0960	3.676	.0570	3.811	.1123
x	3.921	.0734	3.689	.0448	3.461	.0377	3.500	.0547
		(66.1%)		(46.7%)		(66.1%)		(48.7%)
βz	0.091	.0023	0.061	.0018	0.055	.0009	0.094	.0032
		(2.1%)		(1.9%)		(1.6%)		(2.8%)
$(1-\beta)\zeta$	0.194	.0254	0.172	.0288	0.160	.0139	0.217	.0347
		(22.9%)		(30.0%)		(24.3%)		(30.9%)
$2\text{cov}(x, \beta z)$		.0000		.0010		-.0012		.0006
$2\text{cov}(x, (1-\beta)\zeta)$		.0026		.0100		.0010		.0084
$2\text{cov}(\beta z, (1-\beta)\zeta)$		.0072		.0094		.0047		.0106

and Søvsten (2023), Lamadon, Lise, Meghir, and Robin (2024), we find that the resetting friction increases the contribution of the current match.

Finally, in Table 10, we keep only the normal hiring wages ($z < z'$). Now, the contribution of the current surplus ($m = \beta z$) is negligible. This is because of our very low estimate of β. Normal job-to-job moves will not generate a big increase in the match surplus. So, the historical residual $u = (1-\beta)\zeta$ dominates βz.

7 Conclusion

In this paper, we develop a sequential auction model with the additional friction that, with some probability, the second bidder wins the auction. We find evidence that maybe one-third of all wage negotiations are of this inefficient type. Moreover, we model the match output as the sum of a part reflecting worker-firm complementarities and a random match-specific component. We find that complementarities are negligible and that the match effect explains around 20% of the total wage variance. We also find evidence that estimating a two-way fixed-effects regression instead inflates the share of the variance that is explained by the worker type and generates a negative sorting covariance. There are many good reasons to believe that the same job might be good for one worker and bad for another, in ways that cannot be described in terms of an interaction between fixed worker and firm characteristics. Our work opens a new avenue for research on how best to describe matching in labor markets.

References

Abowd, J. M., R. H. Creecy, and F. Kramarz (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–334.

Abowd, J. M., K. L. McKinney, and I. M. Schmutte (2019): "Modeling Endogenous Mobility in Earnings Determination," *Journal of Business & Economic Statistics*, 37(3), 405–418.

Abowd, J. M., B. E. Stephens, L. Vilhuber, F. Andersson, K. L. McKinney, M. Roemer, and S. Woodcock (2009): "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in *Producer Dynamics: New Evidence from Micro Data*, NBER Chapters, pp. 149–230. National Bureau of Economic Research, Inc.

Altonji, J. G., A. A. Smith Jr., and I. Vidangos (2013): "Modeling Earnings Dynamics," *Econometrica*, 81(4), 1395–1454.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3), 673–697.

Arellano, M., R. Blundell, and S. Bonhomme (2017): "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 85(3), 693–734.

Azkarate-Askasua, M., and M. Zerecero (2023): "Correcting Small Sample Bias in Linear Models with Many Covariates," Discussion paper, TSE.

Babet, D., O. Godechot, and M. G. Palladino (2022): "In the Land of AKM: Explaining the Dynamics of Wage Inequality in France," working paper or preprint.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014): “Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics,” *American Economic Review*, 104(6), 1551–1596.

Bagger, J., and R. Lentz (2019): “An Empirical Model of Wage Dispersion with Sorting,” *The Review of Economic Studies*, 86(1), 153–190.

Barlevy, G. (2008): “Identification of Search Models using Record Statistics,” *The Review of Economic Studies*, 75(1), 29–64.

Bilal, A., N. Engbom, S. Mongey, and G. L. Violante (2022): “Firm and Worker Dynamics in a Frictional Labor Market,” *Econometrica*, 90(4), 1425–1462.

Boerma, J., A. Tsyvinski, R. Wang, and Z. Zhang (2023): “Composite Sorting,” Working Paper 31656, National Bureau of Economic Research.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): “A Distributional Framework for Matched Employer Employee Data,” *Econometrica*, 87(3), 699–739.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): “Wage Bargaining with On-the-Job Search: Theory and Evidence,” *Econometrica*, 74(2), 323–364.

Camirand Lemyre, Félix, R. J. C., and A. Delaigle (2022): “Semiparametric Estimation of the Distribution of Episodically Consumed Foods Measured With Error,” *Journal of the American Statistical Association*, 117(537), 469–481, PMID: 36091664.

Card, D., J. Heining, and P. Kline (2013): “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *The Quarterly Journal of Economics*, 128(3), 967–1015.

Chen, S., G. B. Dahl, and S. K. and (2005): “Nonparametric Identification and Estimation of a Censored Location-Scale Regression Model,” *Journal of the American Statistical Association*, 100(469), 212–221.

Cruz, M., E. Milet, M. Olarreaga, and J.-M. Solleder (2024): “Labour mobility costs across sectors and regions in a cross-section: How old is the captain?,” CEPR Discussion Papers 19209, C.E.P.R. Discussion Papers.

de Melo, R. L. (2018): “Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence,” *Journal of Political Economy*, 126(1), 313–346.

Dey, M. S., and C. J. Flinn (2005): “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica*, 73(2), 571–627.

Di Addario, S., P. Kline, R. Saggio, and M. Søvsten (2023): "It ain't where you're from, it's where you're at: Hiring origins, firm heterogeneity, and wages," *Journal of Econometrics*, 233(2), 340–374.

Gautier, P. A., and C. N. Teulings (2015): "Sorting And The Output Loss Due To Search Frictions," *Journal of the European Economic Association*, 13(6), 1136–1166.

Gruetter, M., and R. Lalive (2009): "The importance of firms in wage determination," *Labour Economics*, 16(2), 149 – 160.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85, 29–65.

Iranzo, S., F. Schivardi, and E. Tosetti (2008): "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data," *Journal of Labor Economics*, 26(2), 247–285.

Jinkins, D., and A. Morin (2018): "Job-to-job transitions, sorting, and wage growth," *Labour Economics*, 55, 300–327.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US," *European Economic Review*, 50(4), 877–907.

Kline, P., R. Saggio, and M. Solvsten (2020): "Leave-out estimation of variance components," *Econometrica*, 88(5), 1859–1898.

Krusell, P., and A. A. Smith, Jr. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Lamadon, T., J. Lise, C. Meghir, and J.-M. Robin (2024): "Labor Market Matching, Wages, and Amenities," Working Paper 32687, National Bureau of Economic Research.

Lentz, R., S. Piyapromdee, and J.-M. Robin (2023): "The Anatomy of Sorting,Evidence From Danish Data," *Econometrica*, 91(6), 2409–2455.

Lewbel, A. (2014): "An Overview of the Special Regressor Method," in *The Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*. Oxford University Press.

Lewbel, A., and O. Linton (2002): "Nonparametric Censored and Truncated Regression," *Econometrica*, 70(2), 765–779.

Lindenlaub, I., and F. Postel-Vinay (2023): "Multidimensional Sorting under Random Search," *Journal of Political Economy*, 131(12), 3497–3539.

Lise, J., C. Meghir, and J.-M. Robin (2016): "Matching, sorting and wages," *Review of Economic Dynamics*, 19, 63 – 87, Special Issue in Honor of Dale Mortensen.

Lise, J., and F. Postel-Vinay (2020): "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 110(8), 2328–2376.

Lise, J., and J.-M. Robin (2017): "The Macrodynamics of Sorting between Workers and Firms," *American Economic Review*, 107(4), 1104–35.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

Moffitt, R., and P. Gottschalk (2002): "Trends in the Transitory Variance of Earnings in the United States," *Economic Journal*, 112(478), C68–C73.

OECD (2024): *Promoting Better Career Choices for Longer Working Lives: Stepping Up Not Stepping Out, Ageing and Employment Policies*. OECD Publishing, Paris.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. von Wachter (2018): "Firming Up Inequality*," *The Quarterly Journal of Economics*, 134(1), 1–50.

Sørensen, K. L., and R. Vejlin (2011): "Worker and Firm Heterogeneity in Wage Growth: An AKM Approach," *LABOUR*, 25(4), 485–507.

Sørensen, T., and R. Vejlin (2013): "The importance of worker, firm and match effects in the formation of wages," *Empirical Economics*, 45(1), 435–464.

Sorkin, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133(3), 1331–1393.

Taber, C., and R. Vejlin (2020): "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88(3), 1031–1069.

Tjaden, V., and F. Wellschmied (2014): "Quantifying the Contribution of Search to Wage Inequality," *American Economic Journal: Macroeconomics*, 6(1), 134–61.

Woodcock, S. D. (2015): "Match effects," *Research in Economics*, 69(1), 100–121.

APPENDIX

A Identification

Bargaining power parameter β

Consider two consecutive employment spells with a normal firm-to-firm transition. We know that the wage resulting from the auction is the second price by checking that wages show a progression in the second spell. Let

$$\begin{aligned} w_1 &= x + \beta z_1 + (1-\beta)\zeta_1 \\ &= x + z_1 - (1-\beta)(z_1 - \zeta_1), \end{aligned}$$

with $0 \leq \zeta_1 \leq z_1$, be a wage from the first spell. Let

$$\begin{aligned} w_2 &= x + \beta z_2 + (1-\beta)\zeta_2 \\ &= x + z_2 + (1-\beta)(z_2 - \zeta_2), \end{aligned}$$

with $z_1 \leq \zeta_2 \leq z_2$, be a wage from the second spell.

The hiring wage (the starting wage of the second spell) is

$$\begin{aligned} \underline{w}_2 &= x + \beta z_2 + (1-\beta) z_1 \\ &= \beta(x + z_2) + (1-\beta)(x + z_1). \end{aligned}$$

It is the outcome of Nash bargaining between the worker and the second firm with the first firm's value as threat point. Using the first two wage equations, we can write

$$\underline{w}_2 = \beta \left[w_2 + (1-\beta)(z_2 - \zeta_2) \right] + (1-\beta) \left[w_1 + (1-\beta)(z_1 - \zeta_1) \right].$$

It makes sense to use the last wages observed in each spell for w_1 and w_2 (say $\overline{w}_1, \overline{w}_2$) because the two corresponding worker surpluses $\overline{\zeta}_1, \overline{\zeta}_2$ narrow the wedges $z_1 - \zeta_1, z_2 - \zeta_2$ to their minimal values. In particular, if $\overline{\zeta}_1 = z_1$ and $\overline{\zeta}_2 = z_2$, these wedges are nil and we can solve for β as

$$\beta = \frac{\underline{w}_2 - \overline{w}_1}{\overline{w}_2 - \overline{w}_1}.$$

In the general case,

$$\frac{\underline{w}_2 - \overline{w}_1}{\overline{w}_2 - \overline{w}_1} = \beta \frac{1 + \frac{1-\beta}{\beta} \frac{z_1 - \overline{\zeta}_1}{z_2 - z_1}}{1 - (1-\beta)\frac{z_2 - \overline{\zeta}_2 - (z_1 - \overline{\zeta}_1)}{z_2 - z_1}}$$

remains informative on β as long as $\overline{\zeta}_1$ and $\overline{\zeta}_2$ are close enough to z_1 and z_2.

In practice, by selecting first and second spells with long last steps, we should select trajectories such that $\overline{\zeta}_1$ and $\overline{\zeta}_2$ are close enough to z_1 and z_2 to approximate β well. As an empirical check, we can plot $\frac{\underline{w}_2 - \overline{w}_1}{\overline{w}_2 - \overline{w}_1}$ versus spell durations.

Worker ability x

Now let us consider a spell exhibiting an increasing sequence of wages. Suppose also that it follows a transition from non-employment. So, the first wage of this spell is

$$\begin{aligned}\underline{w}_1 &= x + \beta z_1 \\ &= \beta(x + z_1) + (1-\beta)x \\ &= \beta\left[\overline{w}_1 + (1-\beta)(z_1 - \overline{\zeta}_1)\right] + (1-\beta)x,\end{aligned}$$

where $\overline{w}_1$ is the last wage of this spell and $\overline{\zeta}_1$ is the corresponding worker surplus. If $\overline{\zeta}_1 = z_1$ then x is identified. Otherwise, suppose that a second spell follows that is also increasing. The transition between these two spells must be subject to a normal negotiation yielding the wage

$$\underline{w}_2 = \beta\left[\overline{w}_2 + (1-\beta)(z_2 - \overline{\zeta}_2)\right] + (1-\beta)\left[\overline{w}_1 + (1-\beta)(z_1 - \overline{\zeta}_1)\right].$$

If $\overline{\zeta}_2 = z_2$ then x is identified. And so on.

Increasing sequences starting from non-employment and ending with a best advancement exist with a positive probability. Hence, x is identified if we can observe individual trajectories over a sufficiently long period of time.

In practice, we observe many workers who avoid non-employment throughout the entire observation period. Moreover, stock sampling make the first wage observed (we still call it $\underline{w}_1$) like any other wage, with general form:

$$\underline{w}_1 = x + \beta z_1 + (1-\beta)\underline{\zeta}_1,$$

where $\underline{\zeta}_1$ is the worker surplus indexing the first wage observation. Even if increasing wage trajectories and best advancements occur with a high probability, it will be impossible to disentangle worker ability x from initial worker surplus $\underline{\zeta}_1$. Thus, for many workers we can only compute a posterior distribution of x, for a given parametric specification of the model, which allows us to calculate the likelihood of any individual sequence of observations.

Sequence of match surpluses

Consider a firm-to-firm transition. The hiring wage is of the form

$$\underline{w}_2 = x + \begin{cases} \beta z_2 + (1-\beta) z_1 & \text{if } z_2 > z_1, \\ z_2 & \text{if } z_2 \leq z_1. \end{cases}$$

Suppose that x and z_1 are known. Then, we can deduce the regime of the job-to-job change: $z_2 > z_1$ if and only if $\underline{w}_2 > x + z_1$. Hence, we can deduce z_2 in both regimes. By induction, we can therefore deduce the match productivity for all subsequent employment spells. This is the advantage of the sequential auction mechanism, even imperfect as in the current setup. Employment transitions reset the dependence on past wages (or workers' surplus shares).

Now, we know x and z_1 only if one of the observed past spells is non-employment, or if $\overline{w}_1 = x + z_1$.

Transition rates

Suppose that we have identified all β, x for all workers and z for all matches. The instantaneous probability of leaving non-employment is $\lambda_0 \overline{G}_0(x, 0)$. If some workers accept any job (i.e. there exists x such that $\overline{G}_0(x, 0) = 1$), we can identify λ_0 as the maximum instantaneous probability of leaving non-employment.

Similarly, all wage increases to the maximal wage $\overline{w} = x + z$ occur with probability $\lambda_3 \overline{G}_3(x, z)$. From transitions to a worse state ($z' \leq z$), we identify $\lambda_2 G_2(x, z)$. From transitions to a better match ($z' > z$), we identify $\lambda_1 \overline{G}_1(x, z) + \lambda_2 \overline{G}_2(x, z)$. The layoff rate is $\delta + \lambda_2 G_2(x, 0)$. Hence, we will identify λ_3 if there exists (x, z) such that $\overline{G}_3(x, z) = 1$ and similarly for the other rates.

This requires that $f(x, y)$ be bounded when y varies, and the supports of distributions H_i must also be bounded. For example, $\overline{G}_3(x, z) = 1$ if $z = f(x, \underline{y}) + \underline{v}$ with

$$f(x, \underline{y}) = \min_{y \sim q(y|x)} f(x, y) \text{ and } \underline{v} = \inf \operatorname{Supp}(H_3).$$

Intuitively, we will identify λ_3 from the maximal frequency of job spells such that the last wage lasts unusually longer (over spell characteristics x, z). The rate λ_2 should be identified by the frequency of displacements, or firm-to-firm moves with a fall in wages. The rate λ_1 should be identified by how much more frequent firm-to-firm moves with a wage increase are than moves with a wage cut, and how much more frequent the first wage raises in a job are compared to a maximum promotion. The exogenous layoff rate δ is identified by difference after taking out displacement to non-employment.

Disentangling surplus components

The match surplus in any employment spell out of non-employment is $z = \max\{f(x,y)+v, 0\}$ where v is drawn independently from x, y. This is a standard censored regression, which is identified under the assumption that v has a zero median (see Chen, Dahl, and and, 2005, Lewbel and Linton, 2002, Lewbel, 2014).

If x is a latent variable, this is a much more complicated problem. Camirand Lemyre and Delaigle (2022) recently considered a similar, but different measurement error problem with selection. With our notations, and assuming that the measurement error is linear, their model is observe $z = F + v$ when $z > 0$ with $v \perp\!\!\!\perp F \mid z > 0$. They study this problem under a parametric restriction on $\mathbb{P}(z > 0 \mid F)$. As far as we are aware, there is no solution to the fully nonparametric identification problem.

B Steady state distributions

In this section we derive the steady-state distributions.

First, we define the following integrals.

B.1 Useful integrals

1. Define

$$I_0(t,a,b) = \int_0^t e^{-at'} e^{-b(t-t')}\, \mathrm{d}t' = \begin{cases} \frac{e^{-at}-e^{-bt}}{b-a} & \text{if } a \neq b, \\ te^{-bt} & \text{if } a = b. \end{cases}$$

   Note that $I_0(t,a,b) = I_0(t,b,a)$.

   It follows that

$$\partial_a I_0(t,a,b) = \frac{-te^{-at}}{b-a} + \frac{e^{-at}-e^{-bt}}{(b-a)^2} = \frac{I_0(t,a,b) - I_0(t,a,a)}{b-a},$$
$$\partial_b I_0(t,a,b) = \frac{te^{-bt}}{b-a} - \frac{e^{-at}-e^{-bt}}{(b-a)^2} = \frac{I_0(t,b,b) - I_0(t,a,b)}{b-a}.$$

2. Define, for $a \neq b \neq c$,

$$\begin{aligned} I_1(t,a,b,c) &= \int_0^t I_0\left(t',a,b\right) e^{-c(t-t')}\,\mathrm{d}t' \\ &= \int_0^t \frac{e^{-at'}-e^{-bt'}}{b-a} e^{-c(t-t')}\,\mathrm{d}t' \\ &= \frac{1}{b-a}\left[I_0(t,a,c)-I_0(t,b,c)\right]. \end{aligned}$$

Hence,

$$\partial_a I_1(t,a,b,c) = \frac{\partial_a I_0(t,a,c)+I_1(t,a,b,c)}{b-a},$$

and

$$\partial_b I_1(t,a,b,c) = -\frac{\partial_a I_0(t,b,c)+I_1(t,a,b,c)}{b-a},$$

and

$$\partial_c I_1(t,a,b,c) = \frac{\partial_b I_0(t,a,c)-\partial_b I_0(t,b,c)}{b-a}.$$

B.2 All workers

Let $\ell_0(x)$ denote the distribution of new-born workers and let ξ be the retirement rate. The distribution of individuals by age and type is $\ell(t,x)$ such that

$$\frac{\partial \ell(t,x)}{\partial t} = -\xi\,\ell(t,x).$$

Hence,

$$\ell(t,x) = \ell_0(x)e^{-\xi t}. \tag{4}$$

Note that, for all d,

$$\int_0^t \ell(t',x)e^{-d(t-t')}\,\mathrm{d}t' = I_0\left(t,\xi,d\right)\ell_0(x).$$

B.3 Non-employed workers

Let $u(t,x)$ denote the equilibrium measure of workers (t,x) who are currently not employed.

The law of motion of $u(t,x)$ is

$$\frac{\partial u(t,x)}{\partial t} = [\delta + \lambda_2 G_2(x,0)]\,[\ell(t,x) - u(t,x)] - \left[\xi + \lambda_0 \overline{G}_0(x,0)\right] u(t,x)$$
$$= [\delta + \lambda_2 G_2(x,0)]\,\ell(t,x) - d_0(x) u(t,x),$$

with

$$d_0(x) = \xi + \delta + \lambda_0 \overline{G}_0(x,0) + \lambda_2 G_2(x,0).$$

Jobs terminate either with straight probability δ because of layoff, or because of a displacement shock with a negative surplus draw. Unemployed workers leave non-employment because they retire or find a job.

The solution to this first-order ODE with initial condition $u(0,x) = \ell_0(x)$ (i.e. assuming all workers start their career out of employment) is the cumulative sum of non-employment inflows in time interval $[0,t]$ who have not yet exited non-employment by t. Specifically,

$$u(t,x) = \ell_0(x) e^{-d_0(x)t} + [\delta + \lambda_2 G_2(x,0)] \int_0^t \ell(t',x)\, e^{-d_0(x)(t-t')}\,\mathrm{d}t'$$
$$= \ell_0(x) e^{-d_0(x)t} + [\delta + \lambda_2 G_2(x,0)]\,\ell_0(x) I_0\left(t, \xi, d_0(x)\right).$$

That is,

$$u(t,x) = \left[a(x) e^{-\xi t} + b(x) e^{-d_0(x)t}\right] \ell_0(x), \tag{5}$$

with

$$a(x) = \frac{\delta + \lambda_2 G_2(x,0)}{d_0(x) - \xi}, \quad b(x) = 1 - a(x) = \frac{\lambda_0 \overline{G}_0(x,0)}{d_0(x) - \xi}.$$

Note that, for all d,

$$\int_0^t u(t',x) e^{-d(t-t')}\,\mathrm{d}t' = [a(x) I_0\left(t, \xi, d\right) + b(x) I_0\left(t, d_0(x), d\right)]\,\ell_0(x),$$

and

$$\int_0^t [\ell(t',x) - u(t',x)]\, e^{-d(t-t')}\,\mathrm{d}t' = b(x)\,[I_0\left(t,\xi,d\right) - I_0\left(t, d_0(x), d\right)]\,\ell_0(x).$$

B.4 Matches

Let $m(t,x,z)$ denote the equilibrium number of matches (t,x,z). Let $M(t,x,z)$ denote the number of workers of type (t,x) employed at a job $0 \le z' \le z$,

$$M(t,x,z) = \int_0^z m(t,x,z')\,\mathrm{d}z',$$

with $M(t,x,z) = 0$ for all $z < 0$ (all matches must yield nonnegative surplus), and $M(0,x,z) = 0$ (workers start unemployed).

The law of motion for $M(t,x,z)$ is

$$\begin{aligned}\frac{\partial M(t,x,z)}{\partial t} &= \lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] u(t,x) \\ &\quad + \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right] \left[\ell(t,x) - u(t,x) - M(t,x,z)\right] \\ &\quad - \left[\xi + \delta + \lambda_2 G_2(x,0) + \lambda_1 \overline{G}_1(x,z) + \lambda_2 \overline{G}_2(x,z)\right] M(t,x,z).\end{aligned}$$

The inflow is made of workers who were not in the stock $M(t,x,z)$ and enter; that is, the unemployed and the displaced workers with match surplus greater than z, who draw a new match surplus in $[0,z]$. The outflow comprises retirees (ξ), laid-off ($\delta+\lambda_2 G_2(x,0)$) and workers who are poached or displaced in jobs with a surplus greater than z ($\lambda_1 \overline{G}_1(x,z) + \lambda_2 \overline{G}_2(x,z)$).

Regrouping terms,

$$\begin{aligned}\frac{\partial M(t,x,z)}{\partial t} &= \left(\lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] - \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right]\right) u(t,x) \\ &\quad + \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right] \ell(t,x) - d_1(x,z) M(t,x,z),\end{aligned}$$

where

$$d_1(x,z) = \xi + \delta + \lambda_1 \overline{G}_1(x,z) + \lambda_2.$$

Hence,

$$\begin{aligned}M(t,x,z) &= \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right] \int_0^t \ell(t',x) e^{-d_1(x,z)(t-t')}\, \mathrm{d}t' \\ &\quad + \left(\lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] - \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right]\right) \int_0^t u(t',x) e^{-d_1(x,z)(t-t')}\, \mathrm{d}t' \\ &= \Big[A(x,z)\, I_0\left(t, \xi, d_1(x,z)\right) + B(x,z)\, I_0\left(t, d_0(x), d_1(x,z)\right)\Big] \ell_0(x), \quad (6)\end{aligned}$$

with

$$\begin{aligned}A(x,z) &= \lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] a(x) + \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right] b(x), \\ B(x,z) &= \left(\lambda_0 \left[\overline{G}_0(x,0) - \overline{G}_0(x,z)\right] - \lambda_2 \left[\overline{G}_2(x,0) - \overline{G}_2(x,z)\right]\right) b(x).\end{aligned}$$

We can then deduce

$$\begin{aligned}
m(t,x,z) &= \partial_z M(t,x,z) \\
&= \Big[\partial_z A(x,z)\, I_0\left(t,\xi,d_1(x,z)\right) + \partial_z B(x,z)\, I_0\left(t,d_0(x),d_1(x,z)\right)\Big]\ell_0(x) \\
&\quad + \partial_z d_1(t,z)\Big[A(x,z)\,\partial_b I_0\left(t,\xi,d_1(x,z)\right) + B(x,z)\,\partial_b I_0\left(t,d_0(x),d_1(x,z)\right)\Big]\ell_0(x),
\end{aligned}$$

with

$$\begin{aligned}
\partial_z d_1(t,z) &= -\lambda_1 g_1(x,z), \\
\partial_z A(x,z) &= \lambda_0 g_0(x,z) a(x) + \lambda_2 g_2(x,z) b(x), \\
\partial_z B(x,z) &= \left[\lambda_0 g_0(x,z) - \lambda_2 g_2(x,z)\right] b(x).
\end{aligned}$$

B.5 Wages

Let $n(t,x,y,z,\zeta)$, for $z \geq \zeta \geq 0$, denote the measure of workers whose current wage is

$$w = \gamma(t) + x + \beta z + (1-\beta)\zeta,$$

and the current job surplus is $z = f(x,y) + v$ for some v. Let also

$$N(t,x,y,z,\zeta) = n(t,x,y,z,0) + \int_0^\zeta n(t,x,y,z,\zeta')d\zeta',$$

where $0 < \zeta < z$, and

$$N(t,x,y,z,z) = N(t,x,y,z,z^-) + n(t,x,y,z,z),$$

where $N(t,x,y,z,z^-) = \lim_{\zeta\uparrow z} N(t,x,y,z,\zeta)$ is the left-limit of $N(t,x,y,z,\zeta)$ at $\zeta = z$.

B.5.1 $n(t,x,y,z,0)$

The law of motion for $n(t,x,y,z,0)$ is

$$\frac{\partial n(t,x,y,z,0)}{\partial t} = e_0(x,y,z)u(t,x) - d_3(x,0)n(t,x,y,z,0),$$

for entry and exit rates

$$
\begin{aligned}
e_0(x,y,z) &= \lambda_0\, h_0\left(z - f(x,y)\right) q_0(y|x), \\
d_3(x,0) &= \xi + \delta + (\lambda_1 + \lambda_3)\overline{G}_1(x,0) + \lambda_2.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
n(t,x,y,z,0) &= e_0(x,y,z) \int_0^t u(t',x) e^{-d_3(x,0)(t-t')}\, \mathrm{d}t' \\
&= e_0(x,y,z)\Big[a(x) I_0\left(t,\xi,d_3(x,0)\right) + b(x) I_0\left(t,d_0(x),d_3(x,0)\right)\Big]\ell_0(x).
\end{aligned}
$$

B.5.2 $N(t,x,y,z,\zeta)$

The law of motion for $N(t,x,y,z,\zeta)$, when $0 < \zeta < z$, is

$$
\frac{\partial N(t,x,y,z,\zeta)}{\partial t} = e_0(x,y,z)u(t,x) + e_1(x,y,z)M(t,x,\zeta) - d_3(x,\zeta)N(t,x,y,z,\zeta),
$$

with

$$
\begin{aligned}
e_1(x,y,z) &= \lambda_1\, h_1\left(z - f(x,y)\right) q_1(y|x) + \lambda_2\, h_2\left(z - f(x,y)\right) q_2(y|x), \\
d_3(x,\zeta) &= \xi + \delta + (\lambda_1 + \lambda_3)\overline{G}_1(x,\zeta) + \lambda_2.
\end{aligned}
$$

Hence,

$$
N(t,x,y,z,\zeta) = \int_0^t \left[e_0(x,y,z)u(t',x) + e_1(x,y,z)M(t',x,\zeta)\right] e^{-d_3(x,\zeta)(t-t')} dt',
$$

with

$$
\int_0^t u(t',x)e^{-d_3(x,\zeta)(t-t')}dt' = [a(x)I_0(t,\xi,d_3(x,\zeta)) + b(x)I_0(t,d_0(x),d_3(x,\zeta))]\ell_0(x),
$$

and

$$
\begin{aligned}
&\int_0^t M(t',x,\zeta)e^{-d_3(x,\zeta)(t-t')}dt' \\
&\quad = \ell_0(x) \int_0^t \Big[A(x,\zeta)\, I_0\left(t',\xi,d_1(x,\zeta)\right) + B(x,\zeta)\, I_0\left(t',d_0(x),d_1(x,\zeta)\right)\Big] e^{-d_3(x,\zeta)(t-t')}dt' \\
&\qquad = \Big[A(x,\zeta) I_1\left(t,\xi,d_1(x,\zeta),d_3(x,\zeta)\right) + B(x,\zeta) I_1\left(t,d_0(x),d_1(x,\zeta),d_3(x,\zeta)\right)\Big]\ell_0(x).
\end{aligned}
$$

B.5.3 $n(t, x, y, z, \zeta)$

The law of motion for $n(t, x, y, z, \zeta)$ is

$$\frac{\partial n(t,x,y,z,\zeta)}{\partial t} = e_1(x,y,z)m(t,x,\zeta)+(\lambda_1+\lambda_3)\, g_1(x,\zeta)\, N(t,x,y,z,\zeta)-d_3(x,\zeta)\, n(t,x,y,z,\zeta).$$

Instead of solving this equation, we can more easily deduce $n(t, x, y, z, \zeta)$ from $N(t, x, y, z, \zeta)$ as

$$n(t,x,y,z,\zeta) = \frac{\partial N(t,x,y,z,\zeta)}{\partial \zeta} = \Big(e_0(x,y,z)C_0 + e_1(x,y,z)C_1\Big)\ell_0(x),$$

where

$$\begin{aligned} C_0 &= \frac{\partial}{\partial \zeta}\left[a(x)I_0\left(t,\xi,d_3(x,\zeta)\right) + b(x)I_0\left(t,d_0(x),d_3(x,\zeta)\right)\right] \\ &= -(\lambda_1+\lambda_3)g_1(x,\zeta)\Big[a(x)\partial_b I_0\left(t,\xi,d_3(x,\zeta)\right) + b(x)\partial_b I_0\left(t,d_0(x),d_3(x,\zeta)\right)\Big] \end{aligned}$$

and

$$\begin{aligned} C_1 &= \frac{\partial}{\partial \zeta}\left[A(x,\zeta)I_1\left(t,\xi,d_1(x,\zeta),d_3(x,\zeta)\right) + B(x,\zeta)I_1\left(t,d_0(x),d_1(x,\zeta),d_3(x,\zeta)\right)\right] \\ &= \partial_z A(x,\zeta)I_1\left(t,\xi,d_1(x,\zeta),d_3(x,\zeta)\right) + \partial_z B(x,\zeta)I_1\left(t,d_0(x),d_1(x,\zeta),d_3(x,\zeta)\right) \\ &\quad - \lambda_1 g_1(x,\zeta)\Big[A(x,\zeta)\partial_b I_1\left(t,\xi,d_1(x,\zeta),d_3(x,\zeta)\right) + B(x,\zeta)\partial_b I_1\left(t,d_0(x),d_1(x,\zeta),d_3(x,\zeta)\right)\Big] \\ &\quad -(\lambda_1+\lambda_3)g_1(x,\zeta)\Big[A(x,\zeta)\partial_c I_1\left(t,\xi,d_1(x,\zeta),d_3(x,\zeta)\right)+B(x,\zeta)\partial_c I_1\left(t,d_0(x),d_1(x,\zeta),d_3(x,\zeta)\right)\Big]. \end{aligned}$$

B.5.4 $n(t, x, y, z, z)$

The law of motion for $n(t, x, y, z, z)$ is

$$\begin{aligned} \frac{\partial n(t,x,y,z,z)}{\partial t} &= e_2(x,y,z)\left[\ell(t,x) - u(t,x) - M(t,x,z)\right] \\ &\qquad + \lambda_3 \overline{G}_1(x,z)N(t,x,y,z,z^-) - d_1(x,z)\, n(t,x,y,z,z), \end{aligned}$$

with

$$\begin{aligned} e_2(x,y,z) &= \lambda_2\, h_2\left(z - f(x,y)\right) q_2(y|x), \\ d_1(x,z) &= \xi + \delta + \lambda_1 \overline{G}_1(x,z) + \lambda_2. \end{aligned}$$

Note that we can omit the inflows

$$e_1(x,y,z)m(t,x,z)\,\mathrm{d}z + (\lambda_1+\lambda_3)\, g_1(x,z)\,\mathrm{d}z\, N(t,x,y,z,z^-)$$

because they are negligible.

Hence,

$$n(t,x,y,z,z) = e_2(x,y,z)\int_0^t \left[\ell(t',x) - u(t',x) - M(t',x,z)\right] e^{-d_1(x,z)(t-t')}\,\mathrm{d}t' + \lambda_3\overline{G}_1(x,z)\int_0^t N(t',x,y,z,z^-)e^{-d_1(x,z)(t-t')}\,\mathrm{d}t',$$

with

$$\int_0^t \left[\ell(t',x) - u(t',x)\right] e^{-d_1(x,z)(t-t')}\,\mathrm{d}t' = b(x)\Big[I_0\left(t,\xi,d_1(x,z)\right) - I_0\left(t,d_0(x),d_1(x,z)\right)\Big]\ell_0(x),$$

and

$$\int_0^t M(t',x,z)e^{-d_1(x,z)(t-t')}\,\mathrm{d}t' = \Big[A(x,z)I_1\left(t,\xi,d_1(x,z),d_1(x,z)\right) + B(x,z)I_1\left(t,d_0(x),d_1(x,z),d_1(x,z)\right)\Big]\ell_0(x),$$

and

$$\int_0^t N(t',x,y,z,z^-)e^{-d_1(x,z)(t-t')}\,\mathrm{d}t' = \ell_0(x)\times \Bigg[e_0(x,y,z)\int_0^t \Big[a(x)I_0(t',\xi,d_3(x,z)) + b(x)I_0(t',d_0(x),d_3(x,z))\Big]e^{-d_1(x,z)(t-t')}\,\mathrm{d}t + e_1(x,y,z)\int_0^t \Big[A(x,z)I_1\left(t',\xi,d_1(x,z),d_3(x,z)\right) + B(x,z)I_1\left(t',d_0(x),d_1(x,z),d_3(x,z)\right)\Big]e^{-d_1(x,z)(t-t')}\,\mathrm{d}t\Bigg]$$

where

$$\int_0^t I_0(t',\xi,d_3(x,z))e^{-d_1(x,z)(t-t')}\,\mathrm{d}t = I_1\left(t,\xi,d_3(x,z),d_1(x,z)\right),$$

$$\int_0^t I_0(t',d_0(x),d_3(x,z))e^{-d_1(x,z)(t-t')}\,\mathrm{d}t = I_1\left(t,d_0(x),d_3(x,z),d_1(x,z)\right),$$

and

$$\int_0^t I_1\left(t', \xi, d_1(x,z), d_3(x,z)\right) e^{-d_1(x,z)(t-t')}\,\mathrm{d}t$$
$$= \int_0^t \frac{I_0(t', \xi, d_3(x,z)) - I_0(t', d_1(x,z), d_3(x,z))}{d_1(x,z) - \xi} e^{-d_1(x,z)(t-t')}\,\mathrm{d}t$$
$$= \frac{I_1(t, \xi, d_3(x,z), d_1(x,z)) - I_1(t, d_1(x,z), d_3(x,z), d_1(x,z))}{d_1(x,z) - \xi}$$

and

$$\int_0^t I_1\left(t', d_0(x), d_1(x,z), d_3(x,z)\right) e^{-d_1(x,z)(t-t')}\,\mathrm{d}t$$
$$= \int_0^t \frac{I_0(t', d_0(x), d_3(x,z)) - I_0(t', d_1(x,z), d_3(x,z))}{d_1(x,z) - d_0(x)} e^{-d_1(x,z)(t-t')}\,\mathrm{d}t$$
$$= \frac{I_1(t, d_0(x), d_3(x,z), d_1(x,z)) - I_1(t, d_1(x,z), d_3(x,z), d_1(x,z))}{d_1(x,z) - d_0(x)}.$$

C Likelihood

An individual observation is a sequence $(t_n, s_n, d_n, y_n, w_n, n = 1, ..., N)$ where $t_n = t_1 - 1 + n$ is age/experience, $s_n = U, E, R$ is the employment status (unemployed, employed or retired), $d_n \in \{0, 1\}$ is the indicator of job mobility, y_n is the employer's type and w_n is the wage in period n (both missing if unemployed). Set by default $z_n = \zeta_n = 0$.

C.1 Contribution of the initial observation

C.1.1 Non-employment

The probability density of the joint event "being of type (t, x) and unemployed" is $u(t, x)/L$, where $L = L_0/\xi$ is the size of the population for a flow of young workers L_0 and a retirement rate ξ.

The contribution to the likelihood of Non-employment as first observation is,

$$\mathbb{L}_U(1) = \frac{u(t_1, x)}{L}.$$

Note that worker ability x is unobserved and will have to be integrated out in the end.

C.1.2 Employment

Consider a worker of type (t, x) employed by a firm of type y at a wage w and with job history $\zeta \geq 0$. The current job surplus is $z = f(x, y) + v$. The observed wage is

$$w = \gamma(t) + x + \beta z + (1 - \beta)\zeta,$$

and three cases can occur regarding job history: $\zeta = 0$, $\zeta \in]0, z[$ or $\zeta = z$. But we do not which of the three regimes is observed.

If $\zeta = 0$, then the wage density is $\frac{1}{\beta}\frac{1}{L}n\left(t, x, y, \frac{w-\gamma(t)-x}{\beta}, 0\right)$ as the job surplus is necessarily $z = \frac{w-\gamma(t)-x}{\beta}$. If $\zeta \in]0, z[$, we can decide to fix $z = \zeta + \frac{w_1-\gamma(t)-x-\zeta}{\beta}$, where $z > \zeta > 0$, that is

$$0 < \zeta < w - \gamma(t) - x.$$

Finally, if $\zeta = z$, the wage density is $\frac{1}{L}n(t, x, y, z, z)$ for $z = w - \gamma(t) - x$. Note the absence of a derivative for the transformation $z \mapsto w$.

The contribution to the likelihood of the initial observation is therefore

$$\begin{aligned} \mathbb{L}_E(1) = &\frac{1}{\beta}\frac{1}{L}n(t_1, x, y_1, z_1, \zeta_1) \times \mathbf{1}\{0 \leq \zeta_1 < w_1 - \gamma(t_1) - x\} \\ &+ \frac{1}{L}n(t_1, x, y_1, z_1, z_1) \times \mathbf{1}\left\{\zeta_1 = w_1 - \gamma(t_1) - x\right\}, \end{aligned}$$

for

$$z_1 = \zeta_1 + \frac{w_1 - \gamma(t_1) - x - \zeta_1}{\beta}.$$

The match type ζ_1 is unobserved and will have to be integrated out in the end together with x. Note that the integration with respect to ζ_1 has to be done taking into account that $\zeta_1 = 0$ and $\zeta_1 = w_1 - \gamma(t_1) - x$ are two mass points given x.

C.2 Contribution of subsequent observations

Consider period $n > 1$. For notational simplicity, we "prime" observations in period n, like $w_n = w'$, and we "un-prime" observations in period $n - 1$, like $w_{n-1} = w$. Also, $t_{n-1} = t - 1$ and $t_n = t$.

C.2.1 Retirement

The probability of retiring ($s_n = R$) from any labor-active state ($s = U, E$) is

$$\mathbb{L}_{sR}(n) = \xi.$$

The probability of remaining inactive is $\mathbb{L}_{RR}(n) = 1$ as retirement is an absorbing state.

C.2.2 Non-employment given Non-employment

The probability of remaining non-employed for another period ($s_n = U$) is the probability of not retiring and not finding an acceptable employment:

$$\mathbb{L}_{UU}(n) = 1 - \xi - \lambda_0 \overline{G}_0(x, 0).$$

Note that we neglect the possibility of leaving non-employment to re-enter before the end of the period.

C.2.3 Employment given non-employment

An unemployed worker draws an offer with probability λ_0; this offer comes from a firm y' with probability $q_0(y'|x)$; the wage is

$$w' = \gamma(t) + x + \beta \left[f(x, y') + v' \right],$$

with density $h_0(v')/\beta$, where y' and v' are subject to the positive-surplus condition $z' = f(x, y') + v' \geq 0$.

Hence, this observation contributes to the likelihood for

$$\mathbb{L}_{UE}(n) = \lambda_0 \, q_0(y_n|x) \frac{1}{\beta} h_0(z_n - f(x, y_n)) \times \mathbf{1}\{w_n \geq \gamma(t_n) + x\},$$

with

$$z_n = \frac{w_n - \gamma(t_n) - x}{\beta} \quad \text{and} \quad \zeta_n = 0.$$

C.2.4 Non-employment given employment

This observation contributes to the likelihood for

$$\mathbb{L}_{EU}(n) = \delta + \lambda_2 G_2(x, 0).$$

Either this is straight layoff (δ) or this is a displacement shock and the z that is drawn is negative.

C.2.5 Employment in the same firm, constant wage

There is no wage adjustment between the two periods (ie $w' - \gamma(t) = w - \gamma(t-1)$) if there is no retirement, layoff and displacement. In case of poaching, any unrestricted poaching by a firm $z' > \zeta$ will make the wage increase or induce mobility. If poaching is interrupted, and $z > \zeta$, i.e. the current wage is less than maximum, then $z' > \zeta$ leads to a wage increase. However, if $\zeta = z$, the current wage is maximum and no λ_3-poaching will do anything.

This observation contributes to the likelihood for

$$\mathbb{L}_{EE}(n) = 1 - \xi - \delta - [\lambda_1 + \lambda_3 \mathbf{1}\{\zeta_{n-1} < z_{n-1}\}]\, \overline{G}_1(x, \zeta_{n-1}) - \lambda_2,$$

and we set

$$\zeta_n = \zeta_{n-1} \quad \text{and} \quad z_n = z_{n-1}.$$

C.2.6 Employment in the same firm, wage increase less than maximum

Let z be the surplus of the job, already inferred from period $n-1$. Let w' be the new wage, such that $\Delta = w' - \gamma(t) - (w - \gamma(t-1)) > 0$. It is of the form

$$w' = \gamma(t) + x + \beta z + (1-\beta)\zeta',$$

where $\zeta' = f(x, y') + v'$ is drawn, given x, with probability density $g_1(x, \zeta')$. Poaching does not induce job-to-job mobility but generates a wage increase to a lower wage than the maximum wage if $z > \zeta' > \zeta$.

This observation contributes to the likelihood for

$$\mathbb{L}_{EE}(n) = (\lambda_1 + \lambda_3) \frac{g_1(x, \zeta_n)}{1-\beta} \times \mathbf{1}\left\{\zeta_n \in]\zeta_{n-1}, z_{n-1}[\right\},$$

with

$$\zeta_n = \frac{w_n - \gamma(t_n) - x - \beta z_n}{1-\beta} \quad \text{and} \quad z_n = z_{n-1}.$$

In practice, we decide that the wage is maximum if

$$\frac{z_n - \zeta_n}{z_n} < \alpha$$

with α small. We have experimented values between 0.01% and 5%, and we use 2.5%.

C.2.7 Employment in the same firm, maximum wage increase

Let z be the surplus of the job, already inferred from period $n-1$ observation. Suppose that the new wage w' is the maximum wage

$$w' = \gamma(t) + x + z.$$

This occurs with probability mass $\lambda_3\overline{G}_1(x,z)$.

This observation contributes to the likelihood for

$$\mathbb{L}_{EEmax}(n) = \lambda_3\overline{G}_1(x, z_{n-1}),$$

and we set

$$\zeta_n = z_n = z_{n-1}.$$

C.2.8 Employment in a different firm, new wage greater than previous maximum

In case of unrestricted poaching (including displacement) the new wage is of the form

$$w' = \gamma(t) + x + \beta z' + (1-\beta)z$$

with $z' = f(x,y') + v' > z$.

The observation (y', w') contributes to the likelihood for

$$\mathbb{L}_{EE'}(n) = \frac{1}{\beta}\left[\lambda_1 q_1(y_n|x)h_1(z_n - f(x,y_n)) + \lambda_2 q_2(y_n|x)h_2(z_n - f(x,y_n))\right] \times \mathbf{1}\{w_n > \gamma(t_n) + x + z_{n-1}\},$$

and

$$z_n = \frac{1}{\beta}\left(w_n - \gamma(t_n) - x - (1-\beta)z_{n-1}\right),\ \zeta_n = z_{n-1}.$$

C.2.9 Employment in a different firm, new wage lower or equal than previous maximum

This case indicates a displacement. The new wage is of the form

$$w' = \gamma(t) + x + z'$$

with $\zeta' = z' = f(x,y') + v' < z$.

The observation (y', w') contributes to the likelihood for

$$\mathbb{L}_{EE'}(n) = \lambda_2 q_2(y_n|x)\, h_2(z_n - f(x, y_n)) \times \mathbf{1}\{w_n \leq \gamma(t_n) + x + z_{n-1}\},$$

and

$$z_n = \zeta_n = w_n - \gamma(t_n) - x.$$

C.3 Individual likelihood

Now we can write the complete likelihood for any worker in the sample as

$$\mathbb{L}(x, \zeta_1) = \mathbb{L}_{s_1} \times ... \times \mathbb{L}_{s_N},$$

where we emphasize the two initial variables x and ζ_1. Ability x is always unobserved and ζ_1 is unobserved when $s_1 = E$ while $\zeta_1 = 0$ if $s_1 = U$. We need to integrate x and ζ_1. On which domain?

To answer this question, let us define a spell as a sequence of consecutive time units such that the state does not change: a duration of non-employment or a duration of employment in a given firm. Let $k = 1, ..., K$ index the spell with $k = 1$ indexing the first employment spell. If the first observed spell is non-employment, we treat this spell as an additional initial spell indexed by $k = 0$. Let $s_k = U, E, R$ indicate whether it is a non-employment or an employment spell. If there is no pre-non-employment spell, s_0 is missing. Note also that retirement can only occur in the last spell. Using d_k for the duration of each spell, we can easily simplify the expression of the likelihood.

Within each employment spell, we know that the net log wage $w - \gamma(t)$ increases by steps. Let us index by $j = 1, ..., J_k$ the different steps within an employment spell. Thus w_{kj} will denote the jth rung in spell k if k is an employment spell ($s_k = E$). Note that $w_{k1} = \underline{w}_k$ is the minimal net log wage and $w_{kJ_k} = \overline{w}_k$ is the maximum one. We can then use similar notations for z_k and ζ_{kj} (with $\zeta_{k1} = \underline{\zeta}_k$ and $\zeta_{kJ_k} = \overline{\zeta}_k$). It will also be useful to calculate the duration d_{kj} of each subspell j.

If the first spell is a non-employment spell and after each new spell out of employment, the variable $\underline{\zeta}_1$ is reset to zero. Only for trajectories which start by an employment spell, and only for the part until a spell out of employment occurs, do we need to integrate $\underline{\zeta}_1$ out. Conditional on $\underline{\zeta}_1$, observations generate constraints on x which are informative on its location, as we now explain.

C.3.1 Bounds for $\underline{\zeta}_1$ given x

Consider the first spell and let us assume it is an employment spell.

One single rung. Suppose that the detrended wage does not increase during the spell:

$$w_1 = \underline{w}_1 = \overline{w}_1 = x + \beta z_1 + (1-\beta)\zeta_1,$$

and the only information that we have is that

$$0 \leq \zeta_1 \leq w_1 - x.$$

The integration with respect to ζ_1 has to be done taking into account that $\zeta_1 = 0$ and $\zeta_1 = w_1 - x$ are two mass points given x:

$$\frac{1}{\beta}\frac{1}{L}n\left(t_1, x, y_1, \frac{w_1 - x}{\beta}, 0\right) + \int_0^{w_1 - x} \frac{1}{\beta}\frac{1}{L}n(t_1, x, y_1, z_1, \zeta_1)\,\mathrm{d}\zeta_1 \\ + n\left(t_1, x, y_1, w_1 - x, w_1 - x\right),$$

where $z_1 = \zeta_1 + \frac{w_1 - x - \zeta_1}{\beta}$. The middle integral corresponds to the case where $z_1 > w_1 - x$ or $\zeta_1 < z_1$.

More rungs. Now suppose that $\overline{w}_1 > \underline{w}_1$. Then, for sure, the initial $\zeta_{11} = \underline{\zeta}_1 < z_1$. All subsequent wages contribute to increase the current value of ζ. For the last and greatest wage, we have

$$\underline{w}_1 = x + \beta z_1 + (1-\beta)\underline{\zeta}_1,$$
$$\overline{w}_1 = x + \beta z_1 + (1-\beta)\overline{\zeta}_1.$$

The last equation tells us that $0 < \overline{\zeta}_1 \leq \overline{w}_1 - x$ (as with one single rung). The two equations imply that

$$\overline{\zeta}_1 - \underline{\zeta}_1 = \frac{\overline{w}_1 - \underline{w}_1}{1-\beta}.$$

We finally obtain the following constraint on $\underline{\zeta}_1$:

$$0 \leq \underline{\zeta}_1 \leq \overline{w}_1 - x - \frac{\overline{w}_1 - \underline{w}_1}{1-\beta}. \tag{7}$$

The integration of the likelihood with respect to $\underline{\zeta}_1$ has to be done taking into account that $\underline{\zeta}_1 = 0$ and $\underline{\zeta}_1 = \underline{\zeta}_1^* := \overline{w}_1 - x - \frac{\overline{w}_1 - \underline{w}_1}{1-\beta}$ are two mass points for $\underline{\zeta}_1$ given x:

$$\frac{1}{\beta}\frac{1}{L}n\left(t_1, x, y_1, \frac{\underline{w}_1 - x}{\beta}, 0\right) + \int_0^{\underline{\zeta}_1^*} \frac{1}{\beta}\frac{1}{L}n(t_1, x, y_1, z_1, \underline{\zeta}_1)\,\mathrm{d}\underline{\zeta}_1 + n\left(t_1, x, y_1, \underline{\zeta}_1^*, \underline{\zeta}_1^*\right),$$

where $z_1 = \underline{\zeta}_1 + \frac{\underline{w}_1 - x - \underline{\zeta}_1}{\beta}$.

C.3.2 Bound for x

For individuals who are unemployed in all periods: the upper bound of the integral for is $+\infty$. For individuals who are employed for at least one period, we will only consider the restriction $\underline{\zeta}_1^* \geq 0$. That is,

$$x \leq \overline{w}_1 - \frac{\overline{w}_1 - \underline{w}_1}{1-\beta}.$$

Note that this constraint applies to all employment spells:

$$x \leq \overline{w}_k - \frac{\overline{w}_k - \underline{w}_k}{1-\beta}.$$

We finally use the minimal bound:

$$x \leq \min_k \overline{w}_k - \frac{\overline{w}_k - \underline{w}_k}{1-\beta}.$$

C.4 Numerical integration

In order to calculate the outer integrals with respect to x, we note that densities $u(t,x)$ and $n(t,x,y,z,\zeta)$ are proportional to $\ell_0(x) \propto e^{-\frac{(x-\mu)^2}{2s}}$. So we have to calculate integrals of the form

$$\frac{1}{s\sqrt{2\pi}}\int_{-\infty}^{a} f(x)e^{-\frac{(x-\mu)^2}{2s}}\,\mathrm{d}x = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{0} f(a+\sqrt{2}st)e^{-\frac{\sqrt{2}(a-\mu)t}{s} - \frac{(a-\mu)^2}{2s^2}} e^{-t^2}\,\mathrm{d}t \equiv \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} g(t)e^{-t^2}\,\mathrm{d}t$$

for

$$g(t) = \begin{cases} \frac{1}{2}f(a+\sqrt{2}st)e^{-\frac{\sqrt{2}(a-\mu)t}{s} - \frac{(a-\mu)^2}{2s^2}} & \text{if } t \leq 0, \\ \frac{1}{2}f(a-\sqrt{2}st)e^{\frac{\sqrt{2}(a-\mu)t}{s} - \frac{(a-\mu)^2}{2s^2}} & \text{if } t > 0. \end{cases}$$

This integral can then be approximated by Gauss-Laguerre or Gauss-Hermite quadrature.

D Model simulation

We simulate an economy given parameters as follows.

1. Given α_2, α_3 predict $y = y_1 + \alpha_2 y_2 + \alpha_3 y_3$ from factors y_1, y_2, y_3 (wage location and spread and firm size). Fit the distribution of y across firms in the data. Specifically, we fit a Gamma distribution shifted on $y - \min(y)$. For example, using the whole sample, we estimate a shape parameter of 9.11, a scale parameter of 0.477. Draw M different values of y from this distribution corresponding to M different firms.

2. Use the model to simulate N individuals for 10 periods. In practice we use $M = 5000$ and $N = 50000$ as in France the ratio of firms to workers is about 1:10 (about 3 mil firms, 30 mil workers). For each individual in each period (including the initial period), we draw a firm type y from the continuous distributions $q(y \mid x)$ that we assign to the firm in the pool with the nearest y. This way we have workers classified into discrete firms and they move across these firms over time. We also get a distribution of firm sizes in each period. The overall distribution of matched firms has been verified to be stable over time, and is similar to the empirical density of firm's y across firms.

3. We finally check that our method for estimating firm heterogeneity works. Using the simulated data, we compute for each firm summary statistics of the wage distribution of their employees (mean, median, p25, p75) and its size, based on the distinct workers that it employs. There are 4651 firms. We keep firms with at least 3 workers (consistent with sample selection procedure in actual data.) There are 4015 firms left. As with the data, we use PCA to obtain predicted factors $\widehat{y}_1, \widehat{y}_2$ and $\widehat{y}_3$ is the simulated log size. We then regress the "true" $y = y_1 + \alpha_2 y_2 + \alpha_3 y_3$ on $\widehat{y}_1, \widehat{y}_2, \widehat{y}_3$. The prediction $\widehat{y}$ from this regression is then plotted against the "true" y. The fit is very good (see Figure D.1).

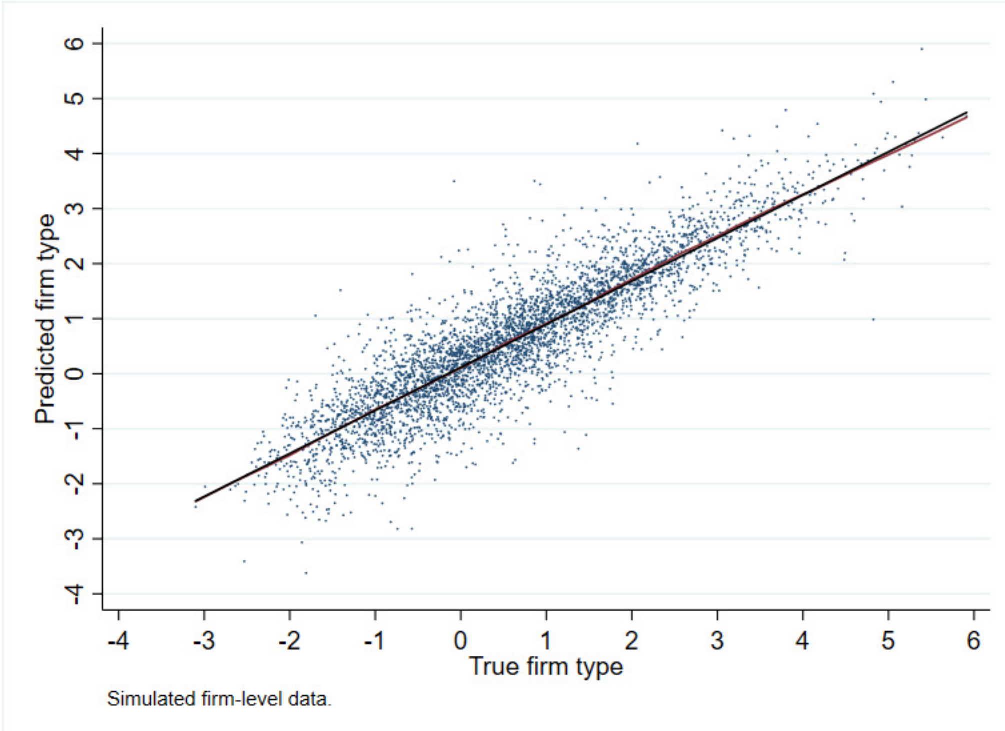


Figure D.1: True versus predicted firm type

E Data Appendix

The DADS panel data records, for each calendar year, all job spells of workers born in October. A unique worker identifier allows to follow workers across years. For each job spell, we observe the starting date (normalized to 1 if the job spell started the previous year), the ending date (normalized to 360 if the job spell ended the following year), and the total earnings received from the job during the year. We start by dividing the 5 years of 2015-2019 into 10 semesters. To do so, we redefine for each semester the starting and ending dates of each job spell according to similar normalizations (e.g. the ending date is 180 for the first semester of a given year if the job goes on the next semester). Other variables are kept unchanged. We then compute the number of paid days in the semester for each job spell from the renormalized starting and end dates. We finally consider as our wage concept the daily wage constructed as the net earnings received during *the whole calendar year* (variable "sn") divided by the number of days worked in this year (variable "dp").[18] Since we will focus on full-time workers (see below), this wage concept effectively corresponds to a wage rate.

For each semester, we then create worker-level data by choosing a unique job spell per worker according to the following priority rules:

[18]This is because there is no way to get semester-level information on daily wage (within a job spell workers may get a pay raise and we do not know when this occurs, neither the magnitude of the raise).

- the job spell corresponds to a job we do not want to drop (see details below),
- the job spell is the one in which the worker spent the largest number of days during the semester,
- the annual earnings are largest (in case there are still multiple eligible job spells after applying the first two rules).

We then recode as non-employed workers those that have worked less than 30 days in their main job spell defined as above and "rectangularize" the data to have a balanced panel with no missing worker*semester observations. Workers are also assumed to be non-employed when they are not observed in the data a given semester.

We apply additional restrictions to this balanced panel. First, we keep prime-age male workers (25-55 y.o. in 2015) that are, whenever employed during the observation period of 2015-2019, always working full-time and always working in a private firm with at least 3 employees (so that a firm type can be computed from the firm-level wage distribution). We further remove workers whose recorded occupation is farmer, CEO, craftsman or merchant at some point in the observation period.

We define the workers' occupation category as the one in which they have been observed employed for the largest number of semesters.

We then apply the following additional cleaning and regularization steps for the whole sample and each skill group separately:

1. Remove individuals having extreme wages at some point of their trajectories: we compute the minimum and maximum observed wage for each worker in the sample and drop workers for which the maximum wage are in the top 1% of the maximum wage distribution and workers for which the minimum wage is in the bottom 2% of the minimum wage distribution. Distributions are computed over all individuals in the sample (one observation per individual).

2. Remove individuals having extreme firm types at some point of their trajectories: we compute for each worker the minimum and maximum firm's first latent characteristic (corresponding to the first principal component obtained from the PCA of the firm wage distribution) observed over the worker's whole trajectory. We then drop workers for which the maximum firm characteristic is in the top 0.5% of the maximum firm characteristic distribution and for which the minimum characteristic is in the bottom 0.5% of the minimum firm characteristic distribution. Distributions are computed over all individuals in the sample.

3. Remove individuals having extreme wage changes within a job spell at some point of their trajectory: we compute differences in log wage across subsequent periods (semesters) within a given job spell for a given worker. We then compute for each worker the minimum and the maximum of these changes within job spell and eventually remove workers for which the minimum change within a job spell is in the bottom 5% of the distribution of these minimum changes and workers for which the maximum change within a job spell is in the top 5% of the distribution of these maximum changes. Distributions are again computed over all individuals in the sample.

4. Keep individuals with initial age (i.e. in period 1) between 25 and 45 y.o.

5. Apply the noise reduction procedure described in Section 4.3.

Table E.1 provides for the whole sample and each skill group the sample characteristics, including an AKM decomposition at each of the steps above.

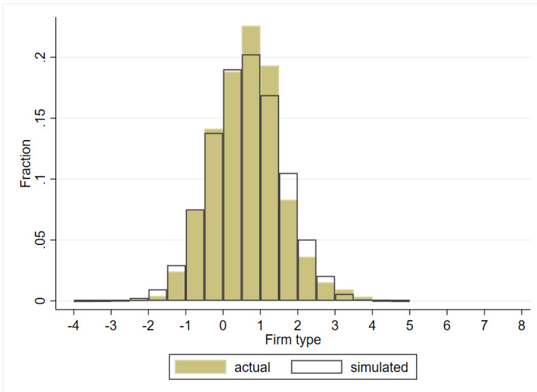
Within the resulting dataset for each skill group, we consider for estimation a random draw of 50,000 individuals (if the data has more).

Table E.1: Descriptive statistics and data cleaning

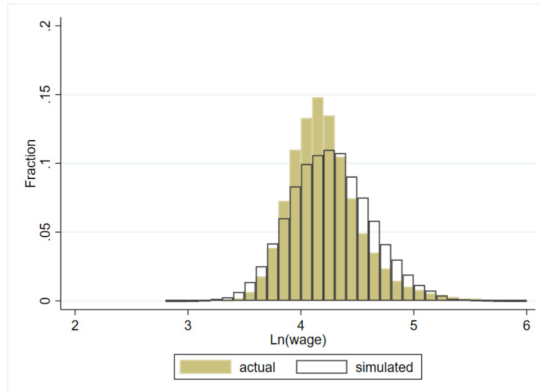
	No obs	Mean	P10	P90	Var	Log wage variance decomposition		
		log wage			log wage	Firm Effect	Wkr Effect	Cov FE/WE
Whole sample								
Initial sample	4 943 118	3.833	3.354	4.422	0.217	0.153	0.916	-0.084
After keeping men	3 370 303	3.864	3.391	4.456	0.217	0.148	0.924	-0.082
After removing extreme wages	3 278 438	3.859	3.401	4.432	0.176	0.149	0.939	-0.083
After removing extreme firm types	3 252 005	3.857	3.402	4.424	0.172	0.149	0.944	-0.085
After removing extreme wage changes	3 004 284	3.858	3.411	4.414	0.164	0.119	0.962	-0.068
After keeping individuals initial age 25-45	2 050 964	3.851	3.421	4.374	0.147	0.130	0.968	-0.074
After wage regularization	2 050 964	3.851	3.427	4.371	0.144	0.128	0.995	-0.077
Managers and engineers								
Initial sample	876 688	4.233	3.793	4.732	0.187	0.012	0.151	-0.016
After keeping men	617 622	4.261	3.822	4.768	0.194	0.012	0.155	-0.026
After removing extreme wages	600 981	4.255	3.833	4.745	0.149	0.009	0.124	-0.020
After removing extreme firm types	595 468	4.253	3.834	4.738	0.146	0.009	0.122	-0.021
After removing extreme wage changes	553 215	4.235	3.832	4.691	0.134	0.006	0.118	-0.015
After keeping individuals initial age 25-45	367 269	4.220	3.843	4.641	0.114	0.007	0.099	-0.006
After wage regularization	367 269	4.220	3.848	4.638	0.112	0.006	0.099	-0.007
Oversees and technicians								
Initial sample	598 360	3.955	3.632	4.299	0.087	0.451	1.235	-0.418
After keeping men	426 295	3.979	3.662	4.320	0.083	0.430	1.229	-0.395
After removing extreme wages	414 568	3.979	3.673	4.309	0.067	0.580	1.369	-0.529
After removing extreme firm types	411 080	3.978	3.675	4.306	0.065	0.569	1.356	-0.517
After removing extreme wage changes	379 584	3.979	3.681	4.299	0.062	0.457	1.312	-0.420
After keeping individuals initial age 25-45	245 081	3.967	3.680	4.277	0.059	0.492	1.332	-0.451
After wage regularization	245 081	3.967	3.685	4.274	0.057	0.495	1.375	-0.459
Admin workers								
Initial sample	628 577	3.675	3.382	4.037	0.096	0.326	0.968	-0.251
After keeping men	206 324	3.720	3.413	4.104	0.098	0.287	0.947	-0.184
After removing extreme wages	200 595	3.718	3.420	4.088	0.078	0.343	0.995	-0.230
After removing extreme firm types	198 908	3.717	3.421	4.082	0.076	0.369	0.997	-0.244
After removing extreme wage changes	183 406	3.721	3.431	4.079	0.071	0.270	1.061	-0.206
After keeping individuals initial age 25-45	135 443	3.719	3.435	4.070	0.070	0.312	1.080	-0.239
After wage regularization	135 443	3.719	3.441	4.067	0.068	0.300	1.094	-0.224
Production workers								
Initial sample	2 503 645	3.773	3.419	4.180	0.123	0.292	0.959	-0.213
After keeping men	1 934 908	3.783	3.442	4.175	0.116	0.295	0.974	-0.220
After removing extreme wages	1 880 749	3.781	3.451	4.158	0.090	0.295	0.952	-0.204
After removing extreme firm types	1 864 990	3.779	3.451	4.154	0.088	0.296	0.951	-0.204
After removing extreme wage changes	1 719 990	3.786	3.464	4.153	0.083	0.246	0.982	-0.175
After keeping individuals initial age 25-45	1 165 910	3.779	3.461	4.140	0.078	0.267	1.017	-0.192
After wage regularization	1 165 910	3.779	3.470	4.135	0.075	0.268	1.066	-0.197

Notes: the initial sample is composed of 25-55 y.o. individuals working always full time in a private firm with at least 3 employees in all their job spells. Log wages correspond to logarithm of daily wages adjusted by a quadratic term in age. We perform an AKM decomposition of log wages by running on the period 2015-2019 the following regression for worker i in firm j at time t: $\ln w_{ijt} = \alpha_i + \beta_j + \gamma_t + \epsilon_{ijt}$ and report in last three columns the contributions of firm effects, worker effects, and their covariance to the variance of log wages.

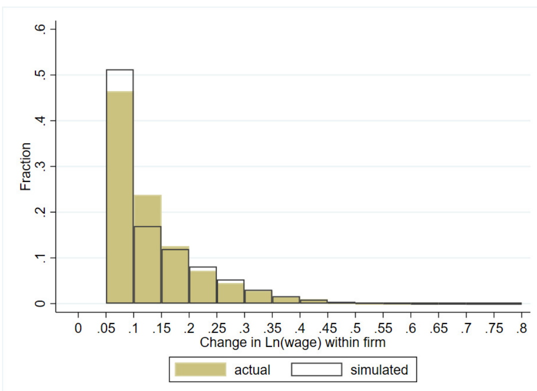
F Additional figures



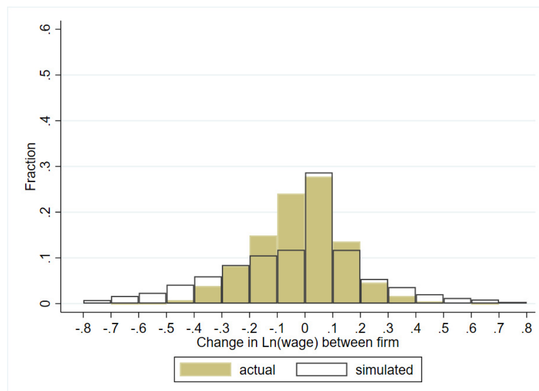
(a) All employer types



(b) All wages

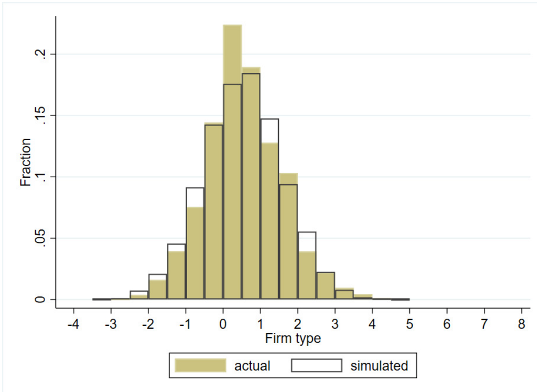


(c) Within-firm wage growth

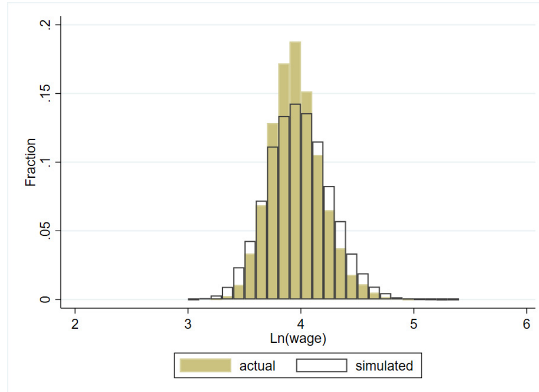


(d) Between-firm wage growth

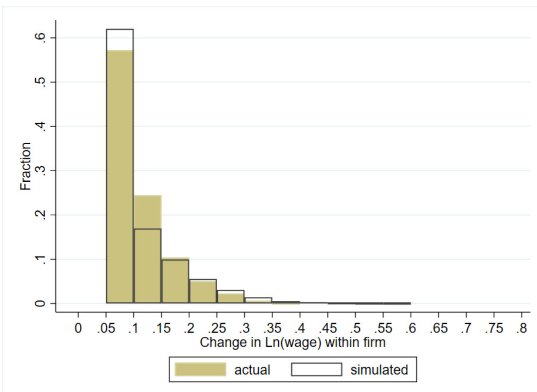
Figure F.2: Fit of wage distributions, Managers



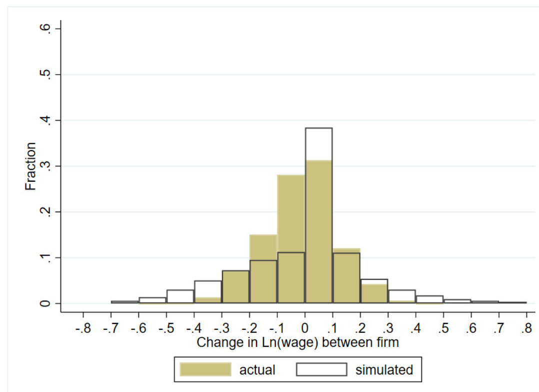
(a) All employer types



(b) All wages

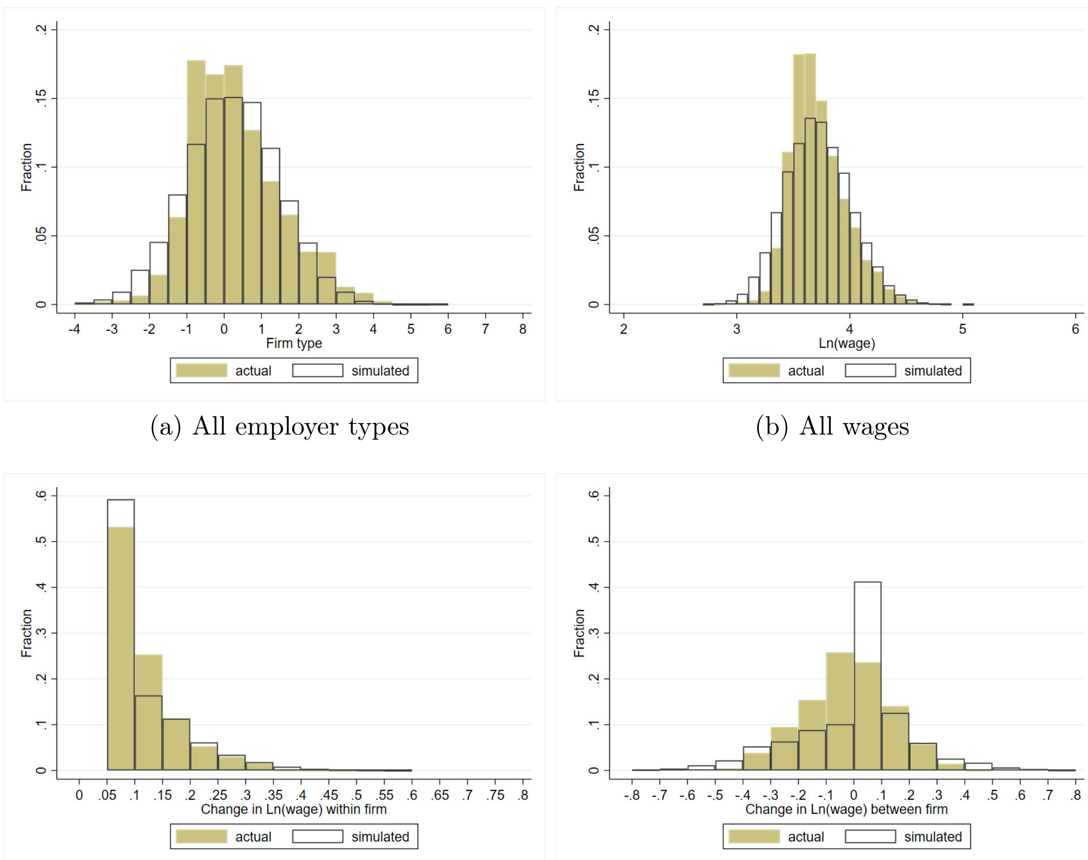


(c) Within-firm wage growth



(d) Between-firm wage growth

Figure F.3: Fit of wage distributions, Supervisors



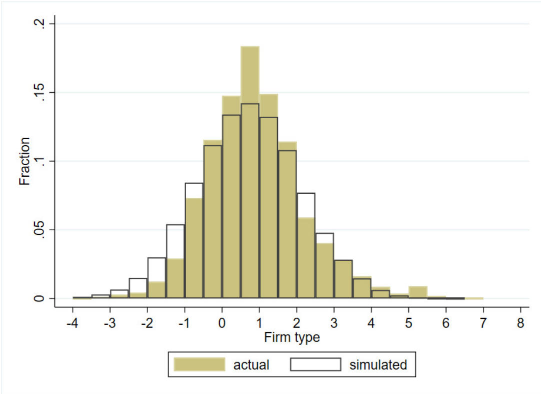
(a) All employer types

(b) All wages

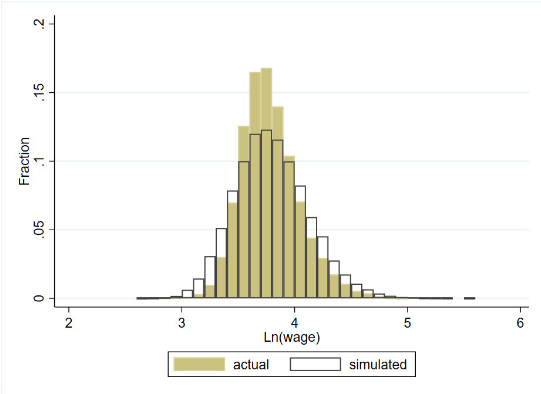
(c) Within-firm wage growth

(d) Between-firm wage growth

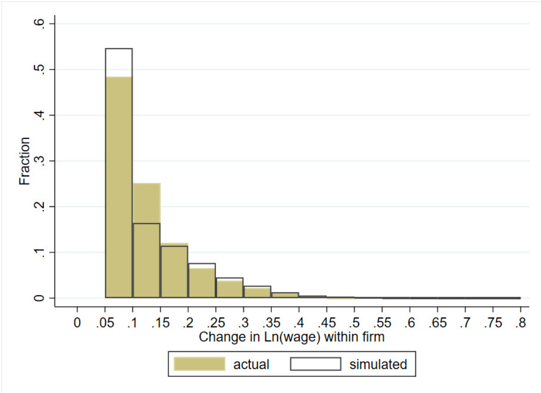
Figure F.4: Fit of wage distributions, Admin



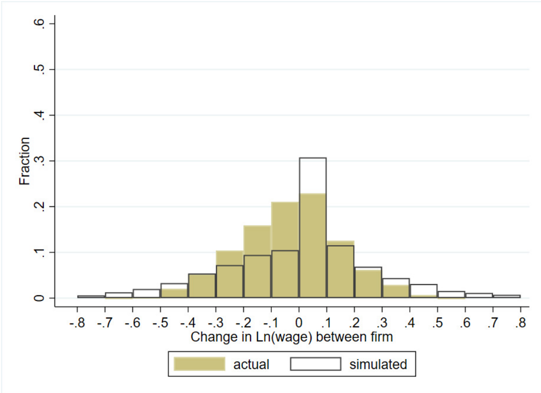
(a) All employer types



(b) All wages

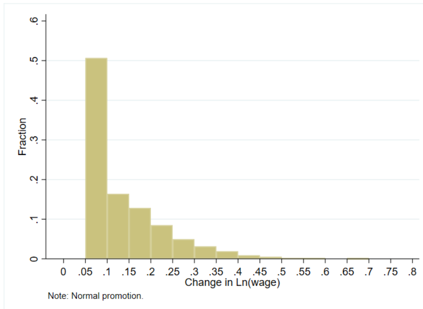


(c) Within-firm wage growth

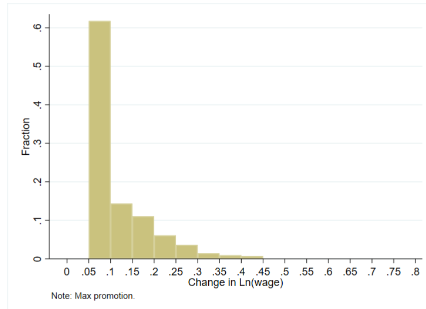


(d) Between-firm wage growth

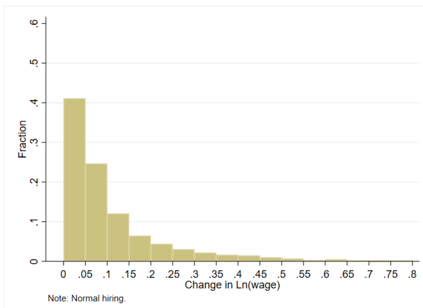
Figure F.5: Fit of wage distributions, Production



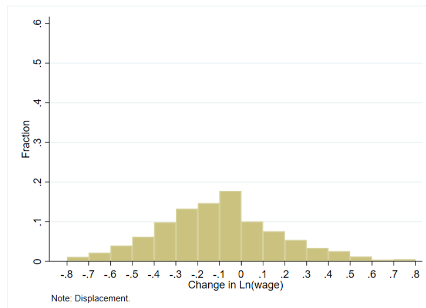
(a) Normal promotions



(b) Max promotions

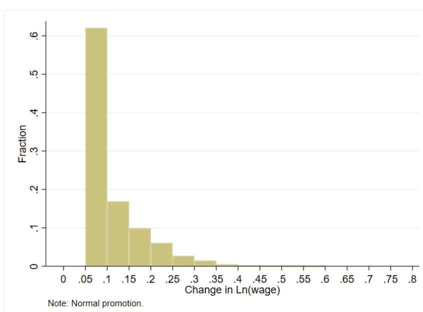


(c) Normal hiring

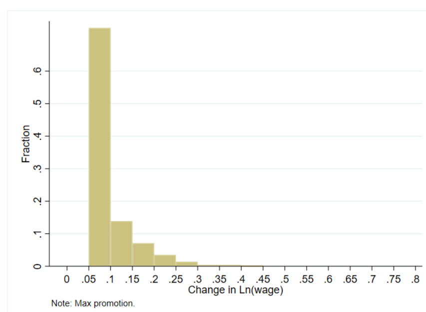


(d) Displacements

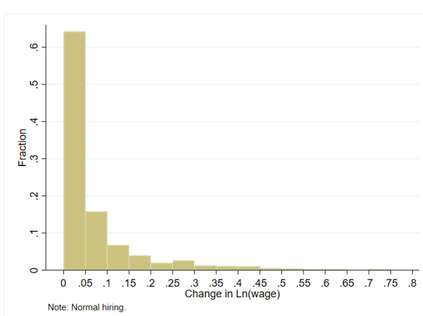
Figure F.6: Simulated wage changes, restricted and unrestricted, Managers



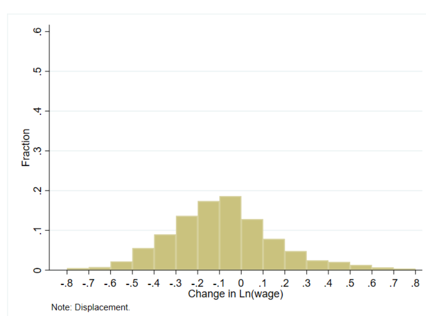
(a) Normal promotions



(b) Max promotions

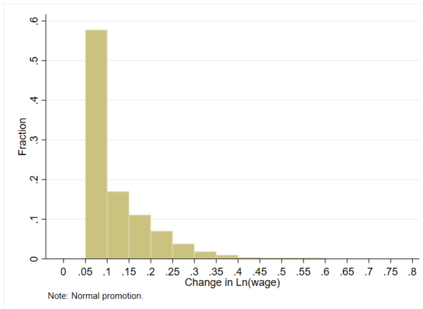


(c) Normal hiring

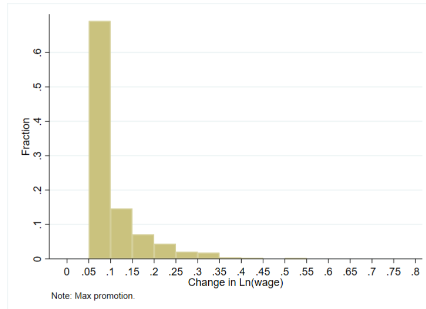


(d) Displacements

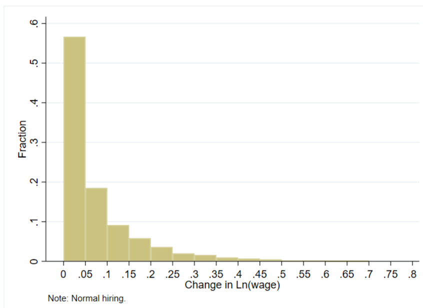
Figure F.7: Simulated wage changes, restricted and unrestricted, Supervisors



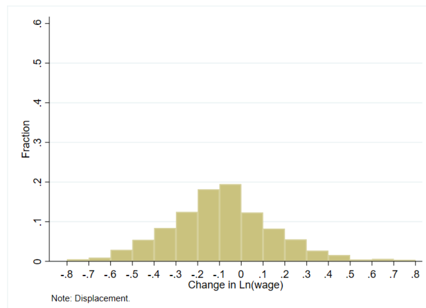
(a) Normal promotions



(b) Max promotions

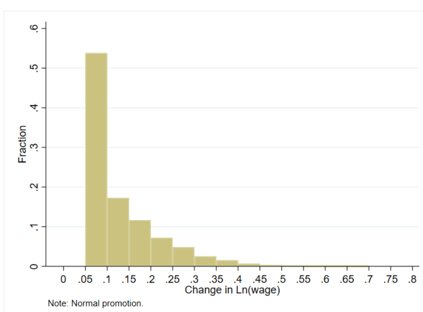


(c) Normal hiring

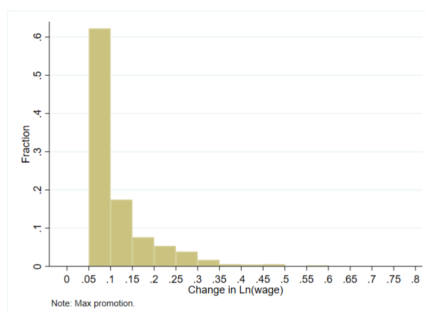


(d) Displacements

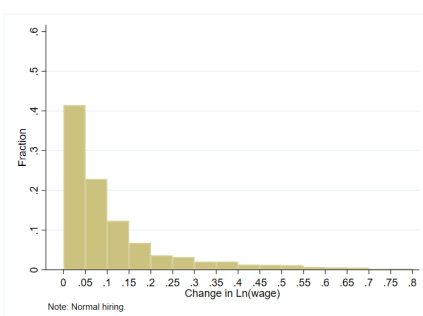
Figure F.8: Simulated wage changes, restricted and unrestricted, Admin



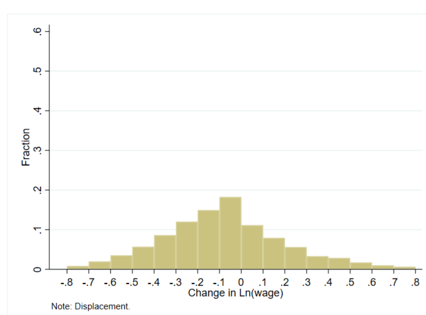
(a) Normal promotions



(b) Max promotions



(c) Normal hiring



(d) Displacements

Figure F.9: Simulated wage changes, restricted and unrestricted, Production