LABOR MARKET MATCHING, WAGES, AND AMENITIES

By

Thibaut Lamadon, Jeremy Lise, Costas Meghir,

and Jean-Marc Robin

July 2024

COWLES FOUNDATION DISCUSSION PAPER NO. 2396

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Labor Market Matching, Wages, and Amenities

Thibaut Lamadon Jeremy Lise Costas Meghir
Jean-Marc Robin*

July 2024

Abstract

This paper develops the nonparametric identification of models with production complementarities, worker-firm specific disutility of labor and search frictions. Mobility in the model is subject to preference shocks, and we assume that firms can write wage contracts. We develop a constructive proof for the nonparametric identification of the model primitives from matched employer-employee data. We use the estimated model to decompose the sources of wage dispersion into worker heterogeneity, compensating differentials, and search frictions that generate between-firm and within-firm dispersion. We find that compensating differentials are substantial on average, but the contribution differs greatly between the lowest and highest types of workers. Finally, we use the model to provide an economic interpretation of several empirical regularities.

*Lamadon: University of Chicago, IFAU, IFS and NBER; Lise: Cornell University, FRB Minneapolis and IFS; Meghir: Yale University, IFAU, IFS and NBER; Robin: Sciences Po. The authors thank our discussant Carlos Carrillo-Tudela and seminar participants at NYU, Wisconsin, Duke, UT Austin, Cornell, DC Search and Matching, FRB Richmond, FRB Minneapolis, and the 2021 NBER Summer Institute for helpful comments and discussions. We are deeply thankful to the IFAU (www.ifau.se) for access to and help with the Swedish administrative data. Costas Meghir thanks the Cowles foundation and the Institution for Social and Policy Studies at Yale University for Financial Support. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Introduction

In his influential lecture series, Mortensen (2003) identifies four key economic forces that contribute to the noticeably large cross-sectional wage inequality observed in most developed economies. The first is the inherent productivity differences between workers. The second is a Rosen (1986) compensating differentials channel where wages offset the varying disutility of labor between employers. The third is the wage dispersion channel caused by market frictions. The fourth is the endogenous matching of workers with firms in the presence of match complementarities as in Becker (1973), Sattinger (1993).

The relative importance of each of these channels has direct implications for the efficiency and redistributive effects of labor policies. For example, if all wage inequality is due to compensating differential, redistributive labor taxation might increase inequality in the economy rather than alleviate it.

Although theoretically well defined, when studying these four channels empirically, we face an immediate barrier even in the presence of detailed matched employer-employee data: the productivities and valuations of workers and jobs are neither directly observed nor exogenously shifted. Workers and firms are forward-looking, and wages and employment decisions reflect a combination of all channels. Quantifying the different forces requires combining a coherent theory with a sound identification strategy.

This paper develops a theoretical model of the labor market that includes all four channels, coupled with an identification result for the primitives from matched data. Our identification approach first uses the nonlinear estimator of Bonhomme, Lamadon, and Manresa (2019, hereafter BLM) to elicit the unobserved heterogeneity of workers and firms and then shows how to recover the primitives of the model from type-specific wages and mobility. The identification suggests an estimation strategy that we use to analyze Swedish administrative records. Our first main empirical finding is that workers with different productivities rank firms differently, leading to both worker heterogeneity and composition differences being important factors in explaining wage dispersion between and within employers. Our second is that non-pecuniary aspects of jobs are a significant contributor to wage dispersion, explaining half of within worker wage dispersion, leaving the other half to the effect of search frictions.

Our theoretical model brings together the theory of complementarities of Becker (1973), the presence of compensating differential of Rosen (1986) and the wage setting in the presence of on the job search of Postel-Vinay and Robin (2002). The economy is made up of heterogeneous workers and firms. When a worker and a firm are matched, the firm collects output, pays a wage to the worker, and the worker endures a disutility of working (equivalently, enjoys an amenity). Both production and disutility are flexible functions of the worker and firm type, allowing for complementarity or substitability between chosen partners. Firms are risk neutral, and workers are risk averse, and both are forward-looking. The meeting between firms and workers is constrained by search frictions. Both employed and unemployed workers can search for jobs and when they meet a vacancy, they draw a mobility preference shock. Upon meeting, firms will offer long-term contracts that price in the mobility shocks.

Solving the equilibrium is more challenging than in the existing literature for two main reasons: first, workers are risk-averse, making the joint value of a match dependent on the wage; and second, there are complementarities and on-the-job search. Yet, by defining the surplus as the highest present value the firm can offer to the worker, we are able to characterize the equilibrium in a very concise way. We first show that the optimal contract is a simple extension of the matching of outside offers of Postel-Vinay and Robin (2002), even when firm types are private information and mobility costs are contractible. Second, we show that our surplus satisfies a surprisingly simple and intuitive equation that extends Postel-Vinay and Robin (2002) and Lise, Meghir, and Robin (2016). Third, we show that the equilibrium wage can be directly expressed as a function of the surplus to the worker. These theoretical results are crucial when taking our framework to the data.

The model produces a range of implications for the wages and mobility patterns of workers. The presence of interactions in match value, through production and disutility, combined with capacity constraints, leads to a Becker (1973) role for sorting in the presence of supermodularity. This insight was first extended to random search in Shimer and Smith (2000) and prevails in our context. Importantly, different workers have different ranking across employers, which generates sorting in equilibrium. This sorting will be imperfect, potentially leading to allocative losses compared to a frictionless environment. This also leads to wage differences between employers for the same worker, as surpluses inherit some of the interactions of production. At the same time, firm-specific amenities allow for an additional channel through which firm wages might differ. As in Rosen (1986), firms with comparatively worse disutility will have to compensate the worker in the form of higher wages at the margin. The importance of the non-pecuniary characteristic of jobs has been documented in several recent papers (Sorkin, 2018; Lamadon, Mogstad, and Setzler, 2022; Lentz, Piyapromdee, and Robin, 2023).

These features are important to keep in mind when considering the decomposition of earnings in matched data pioneered by Abowd, Kramarz, and Margolis (1999, hereafter AKM). The lessons from such a decomposition, as presented in Bonhomme et al. (2023) are that firm premia differences account for between 5% and 15% of the cross-sectional variance in earnings, while the sorting of high-paid workers to high-paying firms accounts for 10% to 20%. Our paper provides a structural interpretation of these numbers by linking sorting and wage premia to productivity, disutility, and the presence of search frictions.[1]

A key distinction in the environment we study is that the worker- and firm-effects from an AKM estimation do not correspond to the actual worker- and firm-types of the model. It has been pointed out by Eeckhout and Kircher (2011) and Hagedorn, Law, and Manovskii (2017) that the non-monotonicity in wages is at odds with the log-additive assumption of the AKM decomposition. Here, we consider the log-additive decomposition as an informative moment of the data, but we use a different approach for the identification and estimation of worker- and firm-types.

Our identification employs a series of steps and results. First, we show that our framework is compatible with the assumptions of the nonlinear estimator of BLM. In contrast to AKM estimation, the types from BLM estimation correspond to the model types. And, conditional on estimated types, the estimated wages, mobility, and allocations from BLM are consistent with the data generating process of the model. This is the first step. Next, we show that wages, allocations, and mobility estimates can be inverted to recover the underlying structural parameters of the model. A key property of the model is that the surplus of the worker can be expressed using wages,

[1] Lamadon, Mogstad, and Setzler (2022) provides a similar structural mapping in a model without search frictions.

inversely weighted by the probability of either a wage increase or a move out of the job, all of which are known for each worker and firm pair from the first step. This property holds more generally than in our particular context (as long as the wage contract is optimal) and allows one to identify surpluses quite generally. The final step is to go from surpluses to production and disutilities. This is made possible by the properties of the contract and by the simple expression of the surplus. We show, for example, how the value of a vacancy can be constructed from wages extracted from BLM estimates once the surplus is known.

Although our identification allows for a flexible mobility shock distribution, there is a very transparent interpretation for the case of logistic distribution. In this case, we can use mobility from the BLM estimation stage to directly recover the match surplus (up to scale)[2], then wages can be used to separate production from disutility (amenities). This is the strategy that we employ in estimation.

We finally use our model to investigate the underlying forces leading to the observed sorting, including the potential role of complementarities in production and amenities. We find that heterogeneity across workers accounts for the majority of wage dispersion we observe in the data; there is substantial positive sorting of workers across jobs; and that just under half of within-worker variation in wages is attributable to compensating differentials. Importantly, the share of within-worker variation due to compensating differentials (i.e., wage variation that is not associated with variation in the worker's utility) is markedly different by worker type. It accounts for less than 10% of the variation in wages for workers of the highest type, but more than 50% for the lowest type. We estimate that the process of sorting workers into firms accounts for 26% of the wage dispersion between workers and accounts for 18% of the overall wage dispersion. Finally, we use the model to interpret several well-documented empirical regularities.

Our paper is most closely related to several recent contributions. The first is Hagedorn et al. (2017), which demonstrates the identification of the equilibrium frictional sorting model of Shimer and Smith (2000) where workers search only from

[2]This uses a revealed preference argument similar to Sorkin (2018), albeit for each type separately. See Arcidiacono et al. (2023) who present a partial equilibrium search model with amenities and preference shocks that exploits insights from dynamic discrete choice estimation.

unemployment. In this model, a worker's wage rank within a firm is equivalent to a worker's productivity rank within a firm, and transitivity across firms allows for a complete ranking of workers based on wages. The second is Sorkin (2018), who uses worker mobility patterns to classify firms according to revealed preference, under the assumption of a common ranking of firms by all workers, in the spirit of the wage posting model of Burdett and Mortensen (1998). Unlike these two papers, we provide identification in a class of models in the spirit of Postel-Vinay and Robin (2002) in which the wage may not order worker types within a firm due to compensating differentials from a variety of sources and where workers do not need to agree on a common ranking of firms due to either production complementarities or worker-firm specific disutility of labor or amenities. Our paper shares many characteristics with Taber and Vejlin (2020). An important difference is that our identification proofs provide a mapping that can be directly adapted to estimate the structure of the model, without the need to repeatedly solve and simulate, providing additional transparency. Lamadon, Mogstad, and Setzler (2022) proposes a frictionless model with several of the features we study including interactions in production and in amenities. They demonstrate the fundamental role of interactions in amenities for understanding the distribution of firm size, wage premia and sorting while making clear that the presence of firm effects could be driven by compensating differential rather than market power. In a dynamic context with search friction, we solve the key idenitifcation problem of separating these channels and demonstrate the empirical importance of amenities in understanding systematic worker flows across jobs, particularly for workers with the lowest average pay. Morchio and Moser (2023) find that interactions in amenities explain a substantial fraction of the gender pay gap in Brazil. As in Balke and Lamadon (2022), we model workers as risk-averse and allow firms to choose optimal contracts. In common with these papers, the empirical analysis uses matched employer-employee data, in our case, from Sweden (see Friedrich et al., 2022).

We proceed as follows. In Section 1 we present the model, which we characterize in Section 2. We prove nonparametric identification in Section 3. In Section 4 we discuss the data and present the estimates. We present the model decompositions of the wages and the interpretation of the empirical regularities in Section 5, followed by the conclusion. The proofs are collected in Appendix A.1, with additional derivations and robustness collected in Online Appendix B.

1 The labor market model

We consider a steady-state economy populated with heterogeneous workers and firms. Time is discrete. Workers and firms are forward-looking, discount the future at a rate r, and do not have a finite horizon.

1.1 Agents and states

Workers have heterogeneous skills indexed by x. Firms are heterogeneous in production technology and are indexed by y. We denote by $\ell(x)$ the measure of type x workers in the population, with the total measure normalized to one. The measure of unemployed workers of type x is $\ell_0(x)$ and the measure of matches of type (x, y) is $\ell_1(x, y)$, with $\ell_0(x) + \int \ell_1(x, y)\,\mathrm{d}y = \ell(x)$. The corresponding total measures are L_0 and $L_1 = 1 - L_0$.

Firms post job vacancies and employ workers. Each firm can employ multiple workers, potentially of different types, where the total firm output is the sum of the output for each match, and hiring and wage contract decisions are made independently between jobs. Let $n(y)$ denote the measure of jobs of type y, which may be vacant or matched to a worker, with the total measure N. We denote by $v(y)$ the measure of job openings of type y and by V the total number of vacancies, with $n(y) = v(y) + \int \ell_1(x, y)\,\mathrm{d}x$.

We assume that $\ell(x)$ and $n(y)$ are exogenous, with $v(y)$, $\ell_0(x)$, and $\ell_1(x, y)$ determined in equilibrium. To simplify the notation, we treat x and y as continuous variables, but they can be discrete and will be in the empirical application.

1.2 Timing of events

At the beginning of each period, each worker can be unemployed or employed. The timing of events during the period and for each employment situation is as follows.

Unemployed workers. When not working, workers enjoy the utility of home production $b(x)$ during the period. At the end of the period, each unemployed worker contacts a job vacancy with probability λ_0. This vacancy, y, is drawn from the cross section of vacancies, with probability density $v(y)/V$. Upon meeting, the workers draw an instantaneous utility cost shock ξ from the distribution G_0 (with negative support and density g_0), which they only experience if they move, in this case accepting the job. We assume that firms make take-it-or-leave-it offers to unemployed workers in a way that we will describe below.

Employed workers. A job produces $f(x, y)$ during the period. Workers receive a wage w that they value at utility flow $u(w)$. They incur an overall flow cost $c(x, y)$ from providing labor in this job, net of any amenity value. We will use the terms disutility of labor and amenity interchangeably. Their flow utility from a job is $u(w) - c(x, y)$. Job costs or amenities are predetermined and set up at the foundation of the firm either as fixed characteristics or because of location.

At the end of the period, the match is exogenously destroyed with probability $\delta(x, y)$. The job becomes vacant and the worker becomes unemployed, remaining as such until the following period. The match continues with probability $\overline{\delta}(x, y) \equiv 1 - \delta(x, y)$. The worker may then contact an alternative vacancy with probability λ_1. This vacancy is also drawn from the cross section of vacancies, described by the probability density $v(y')/V$. At the time of meeting, the worker draws an instantaneous mobility shock ξ from the distribution G_1 (with full support and density g_1). We assume that both firms observe the realized ξ and compete for the worker's services in an auction that we will describe below.

Meeting technology. Let $M(L, V)$ be the number of meetings per unit of time, where L is the effective number of workers searching and V is the total number of vacancies. Specifically, $L = L_0 + \kappa L_1$, where κ is the relative search efficiency of the employed workers. We define the equilibrium meeting rates for unemployed and employed workers by

$$\lambda_0 = \frac{M}{L}, \quad \lambda_1 = \kappa \frac{M}{L}, \quad \text{and} \quad \overline{\lambda}_j = (1 - \lambda_j). \tag{1}$$

Similarly, a vacancy meets an unemployed worker with probability $\lambda_0 L_0/V$, drawing a type x from the distribution $\ell_0(x)/L_0$. It meets an employed worker with probability $\lambda_1 L_1/V$, drawing from the cross-sectional distribution $\ell_1(x, y)/L_1$.

1.3 Poaching

Define $W_0(x)$ as the present value of the unemployed worker and $\Pi_0(y)$ as the value of a vacancy. An employment contract promises a value W to the worker and W must be strictly greater than $W_0(x)$ for employment to be preferred to unemployment. We call the difference $R = W - W_0(x) \geq 0$ the surplus of the worker or the value of the contract.

Define $\Pi_1(x, y, R)$ as the largest present value of profit that a firm can achieve when employing the worker and providing that worker with a surplus of R. Implicit in this definition is a contracting space, which we will define later. The smallest value that satisfies the worker's participation constraint is $R = 0$ and whenever $\Pi_1(x, y, 0)$ is positive, there is value to be shared between the worker and the firm. In this case, we say that the match is viable and denote the highest value of R that satisfies the firm's participation constraint by $S(x, y)$.

In Section 1.4 we will characterize the contract that delivers a promised value R. In this subsection, we first describe how the contract value R is determined in the case of poaching, that is, the competition for a worker between two potential employers. Suppose that a worker of type x employed in a firm of type y is approached by another firm of type y'. Poaching opens up the possibility of making a move, which is associated with a mobility value ξ. We assume that both x and ξ are common knowledge. However, as in standard auctions, we do not assume that the firm types are observable.[3]

Definition 1. *The poaching mechanism.* The incumbent and the poaching firms report their reservation surpluses, say B and B', not necessarily equal to $S(x, y)$ and $S'(x, y)$. The mechanism then makes the following prescriptions:

1. If $B' + \xi > \max\{B, \xi\}$, the poacher wins and must deliver the surplus $\max\{B, \xi\} - \xi$ to the worker, who receives a total of $\max\{B, \xi\}$ after the move.
2. If $B \geq \max\{B' + \xi, \xi\} = B'^{+} + \xi$ (with $B'^{+} = \max\{B', 0\}$), the incumbent wins and must provide a surplus of at least $B'^{+} + \xi$ to the worker.
3. If $\xi > \max\{B, B' + \xi\}$ the worker moves to unemployment and collects ξ.

We focus on strategies where the bid B from the incumbent is credible, that is, even in cases where the incumbent loses for sure, there has to exist a feasible strategy that would deliver B to the worker.

1.4 The firm's problem

The employer faces two separate decision problems. First, it must choose a bid $B(\xi)$ in the event that the worker meets a vacancy and draws a mobility shock ξ. Second, for a given promised value R, it must decide how it will be implemented through wages and separations.

[3] See Online Appendix B.1 for an ascending auction interpretation of this mechanism.

An employment contract can be defined recursively as a current worker surplus R, a wage w for the first period and continuation values at the end of the first period: the status quo value R_0 for a worker who is not approached by another firm, and the retention value $R_1(B', \xi)$ for a worker who is contacted by a firm with a bid B' and a mobility shock ξ. Let F_0 be the distribution of bids, which will be characterized later.

Consider a firm y employing a worker x with a promised surplus R. In calculating its profit $\Pi_1(x, y, R)$, the firm anticipates all future events. In case of poaching at the end of the current period, the firm assumes that the poacher will abide by the mechanism. Therefore, if $B(\xi) < B'^{+} + \xi$ the poacher wins, the incumbent firm receives $\Pi_0(y)$, and the worker collects $W_0(x) + \max\{B(\xi), \xi\}$. If $B(\xi) \geq B'^{+} + \xi$, the current employer wins and the mechanism imposes a minimum payment of $W_0(x) + B'^{+} + \xi$ to the worker. It is a minimum payment because it may be below the promised surplus R_1. For any promised worker surplus $R \leq S$, the firm problem is

$$\begin{aligned}\Pi_1(x, y, R) = \max_{\{w, R_0, R_1(B',\xi), B(\xi)\}} \Big\{ & f(x, y) - w + \frac{\delta(x, y)}{1+r}\Pi_0(y) + \frac{\overline{\delta}(x, y)\overline{\lambda}_1}{1+r}\Pi_1(x, y, R_0) \\ & + \frac{\overline{\delta}(x, y)\lambda_1}{1+r} \iint \Big[\mathbf{1}_{\{B(\xi) \geq B'^{+}+\xi\}} \Pi_1(x, y, R_1(B', \xi)) \\ & + \mathbf{1}_{\{B(\xi) < B'^{+}+\xi\}} \Pi_0(y) \Big] \, \mathrm{d}F_0(B'|x, \xi) \, \mathrm{d}G_1(\xi) \Big\}, \quad (2)\end{aligned}$$

subject to the constraints of promise keeping, the auction, and participation:

$$\begin{aligned}W_0(x) + R = u(w) - c(x, y) + & \frac{\delta(x, y)}{1+r} W_0(x) + \frac{\overline{\delta}(x, y)\overline{\lambda}_1}{1+r} (W_0(x) + R_0) \\ & + \frac{\overline{\delta}(x, y)\lambda_1}{1+r} \iint \Big[\mathbf{1}_{\{B(\xi) \geq B'^{+}+\xi\}} (W_0(x) + R_1(B', \xi)) \\ & + \mathbf{1}_{\{B(\xi) < B'^{+}+\xi\}} (W_0(x) + \max\{B(\xi), \xi\}) \Big] \, \mathrm{d}F_0(B'|x, \xi) \, \mathrm{d}G_1(\xi), \quad \text{(PK)}\end{aligned}$$

$$R_1(B', \xi) \geq B'^{+} + \xi \text{ for all } B'^{+} + \xi < B(\xi) \quad \text{(AC)}$$

$$R_0 \geq 0, \text{ and } \Pi_0(y) \leq \Pi_1(x, y, R'), \text{ for } R' \in \{R_0, R_1, B\}. \quad \text{(C)}$$

Equation (2) is understood as follows. First, the firm collects the flow output $f(x, y)$ and pays the wage w. Then, with probability $\delta(x, y)$ the match separates, in which case the firm recovers a vacancy in the next period, whose value is $\Pi_0(y)$. With probability $\overline{\delta}(x, y)\lambda_1$, the worker draws a firm bidding B' from F_0 and a preference shock ξ from G_1. The firm chooses the bid $B(\xi)$ to submit to the auction for each of these encounters and the worker-surplus $R_1(B', \xi)$ to deliver when it retains the worker. Finally, it chooses R_0 for the case where no external contact is made. Each of these choices is conditional on the state (x, y, R).

When making these choices, the firm is subject to the promise-keeping constraint (PK), which states that the value to the worker of the implemented contract should be equal to the promised value $W_0(x) + R$. The worker collects the flow utility $u(w) - c(x, y)$. Then, with probability $\delta(x, y)$ the worker is laid off, in which case the worker obtains zero surplus in the next period. With probability $\overline{\delta}(x, y)\overline{\lambda}_1$ the worker is neither laid off nor poached and continues with R_0. With probability $\overline{\delta}(x, y)\lambda_1$, the worker is contacted by a firm with bid B' drawn from F_0 and ξ drawn from G_1. If $B(\xi) \geq B'^+ + \xi$ the incumbent wins the auction and pays $R_1(B', \xi)$, which should be at least equal to the value $B'^+ + \xi$ returned by the mechanism (AC); If $B(\xi) < B'^+ + \xi$ the worker moves either to the poacher (for a worker surplus equal to $\max\{B(\xi), \xi\}$) or to unemployment (for a worker surplus ξ). Here, we assume that the poacher pays what is required by the mechanism. Finally, (C) represents the participation constraint of the worker, the limited commitment of the firm, and the credibility of the bid.

Theorem 1. *Assume that the utility function u is bounded, twice continuously differentiable, strictly increasing, concave and has bounded non-zero first and second derivatives.*

1. *The profit value $\Pi_1(x, y, R)$ is bounded, continuous, strictly decreasing and strictly concave in R and differentiable almost everywehere.*
2. *It is a weakly dominant strategy to bid $B(\xi) = S(x, y)$, where $\Pi_1(x, y, S(x, y)) = \Pi_0(y)$.*
3. *The optimal contract is such that the chosen wage $w(x, y, R)$ solves*
$$\frac{1}{u'(w)} = -\frac{\partial \Pi_1(x, y, R)}{\partial R},$$
and the optimal continuation values are $R_0 = R$ and $R_1(B', \xi) = \max\{B'^+ + \xi, R\}$, for $B'^+ + \xi \leq S(x, y)$.

Firms have a reservation value $S(x, y)$ that is, the contract surplus that produces minimal profit $\Pi_0(y)$. The incumbent firm is truthful and bids exactly its reservation value. It will offer a constant value to the worker over time until a meeting with an outside firm requires raising the value to retain the worker. The wage remains constant in the absence of an outside offer and increases whenever it is necessary to counter an outside offer. Following a similar argument, truthfulness is also optimal for the poaching firms.

The annuity value for a firm with a vacancy of type y' is:

$$r\Pi_0(y) = \max_{B_0(x,\xi),B_1(x,\xi)} \frac{\lambda_0 L_0}{V} \iint \mathbf{1}_{\{B_0(x,\xi)+\xi>0\}} \left[\Pi_1(x,y,-\xi) - \Pi_0(y)\right] \frac{\ell_0(x)}{L_0} \,\mathrm{d}G_0(\xi)\,\mathrm{d}x \\ + \lambda_1 \iiint \mathbf{1}_{\{B_1(x,\xi)+\xi>B'\}} \left[\Pi_1(x,y,B'-\xi) - \Pi_0(y)\right] \,\mathrm{d}F_1(B',x|\xi)\,\mathrm{d}G_1(\xi). \quad (3)$$

$F_1(x, B'|\xi)$ is the distribution of workers and retention bids conditional on the mobility shock. When the worker is unemployed, the implicit retention bid B' will be 0. When the worker is employed, the firm is faced with a distribution of retention bids coming from the cross-sectional distribution of employed workers, combined with the reporting strategy of each of these firms. The vacant firm chooses the values to report to the auction, denoted $B_0(x, \xi)$ and $B_1(x, \xi)$. We take the minimum value constraint from the auction as the promised value at the start of the match since it satisfies the worker's participation constraint and, unlike for the incumbent firm, there is no preexisting promised value. Whenever the firm wins an auction, it receives a capital gain of $\Pi_1(x, y, R) - \Pi_0(y)$.

The value of unemployment. We assume that employers have full monopsony power when they meet an unemployed worker, which in our auction mechanism is equivalent to being with a firm with zero surplus. When an unemployed worker of type x meets a vacancy of type y, it draws a mobility cost shock ξ from G_0 with negative support. Employers must compensate the worker for this cost in order to hire. There are two possible cases. If $S(x, y) + \xi > 0$, which means $S(x, y) > 0$, the match is viable. The firm hires the worker, paying $-\xi$ to compensate for the cost of mobility. The worker receives a surplus of zero in total. If $S(s, y) + \xi \leq 0$ no offer is made, the worker remains unemployed and also receives zero surplus. The annuity value of unemployment to a type-x worker is then simply:

$$rW_0(x) = (1 + r)b(x). \quad (4)$$

1.5 Equilibrium

In a stationary truth-telling equilibrium the flows out of and into $\ell_1(x,y)$ are equal:

$$\ell_1(x,y)\Big(\delta(x,y)+\overline{\delta}(x,y)\lambda_1\int\overline{G}_1\big(S(x,y)-S(x,y')\big)v(y')\,\mathrm{d}y'\Big) \tag{5}$$
$$=\lambda_0\ell_0(x)\frac{v(y)}{V}\overline{G}_0(-S(x,y))+\lambda_1\frac{v(y)}{V}\int\overline{G}_1\big(S(x,y')-S(x,y)\big)\overline{\delta}(x,y)\ell_1(x,y')\,\mathrm{d}y'.$$

With truth-telling the distribution of retention bids faced by firms are given by:

$$F_0(B'|x,\xi)=\int\mathbf{1}_{\{S(x,y')\leq B'\}}\frac{v(y')}{V}dy',$$
$$F_1(B',x|\xi)=\int\mathbf{1}_{\{S(x,y')\leq B'\}}\overline{\delta}(x,y')\frac{\ell_1(x,y')}{L_1}dy', \tag{6}$$

which simply states that incumbent firms draw outside bids from the cross-section of vacancies, vacancies draw from the cross-section of employed workers, and all firms bid truthfully.

Definition 2. A *stationary search equilibrium with sequential auctions* is characterized by meeting probabilities λ_0 and λ_1; employment measure $\ell_1(x,y)$; bid distributions $F_0(B'|\xi,x),F_1(B'|\xi,x)$; firm value functions $\Pi_0(y),\Pi_1(x,y,R)$ and their respective policies such that:

1. The meeting probabilities λ_0,λ_1 are consistent with the meeting technology (1).
2. Taking F_0,F_1 as given, $\Pi_1(x,y,R)$ and $\Pi_0(y)$ solve equations (2) and (3), which take into account the mobility decisions of the workers.
3. The policies of $\Pi_1(x,y,R)$ are $\Pi_0(y)$ are truth-telling: for firm y employing worker x, $B(\xi)=S(x,y)$ and for vacancy y, $B_0(x,\xi)=B_1(x,\xi)=S(x,y)$.
4. $\ell_1(x,y)$ satisfies (5) and $F_0(B'|\xi,x),F_1(B'|\xi,x)$ are generated by (6).

2 Properties of the equilibrium

In this section, we present properties of the model that will be useful for identification and estimation in Sections 3 and 4. (See the online appendix B.2 for additional details.)

Equilibrium wage equation. The value function for an employed worker can be inverted at the equilibrium to obtain a wage equation in terms of current surplus R

and maximum surplus $S(x, y)$:

$$
\begin{aligned}
u(w(x,y,R)) = c(x,y) + \frac{r+\delta(x,y)}{1+r} R + \frac{r}{1+r} W_0(x) \\
- \frac{\lambda_1 \bar{\delta}(x,y)}{1+r} \int \Bigg[\int_{R-S(x,y')}^{S(x,y)-S(x,y')} (S(x,y') + \xi - R)\, g(\xi)\, \mathrm{d}\xi \\
+ \int_{S(x,y)-S(x,y')}^{\infty} (\max\{\xi, S(x,y)\} - R)\, g(\xi)\, \mathrm{d}\xi \Bigg] \frac{v(y')}{V}\, \mathrm{d}y'. \quad (7)
\end{aligned}
$$

The wage is increasing in promised surplus R. For a given R, the wage increases in the disutility of labor net of amenities $c(x, y)$ and decreases in the maximum surplus of the worker $S(x, y)$. There are compensating differentials for the current $c(x, y)$ (Rosen, 1986) and for the extent of the potential wage growth (Postel-Vinay and Robin, 2002).

Surplus equation. Since $\Pi_1(x, y, S) = \Pi_0(y)$, we can deduce that the maximum wage in an (x, y) match is

$$
\overline{w}(x,y) = w(x,y,S(x,y)) = f(x,y) - \frac{r}{1+r}\Pi_0(y), \quad (8)
$$

We also know that at this point the worker is receiving maximum surplus, and we can substitute equation (8) and $R = S(x, y)$ into the wage equation (7) to obtain the following equation for the maximum worker surplus $S(x, y)$:

$$
\begin{aligned}
(r + \delta(x,y))\, S(x,y) = (1+r)\left[u(\overline{w}(x,y)) - c(x,y)\right] \\
- rW_0(x) + \lambda_1 \bar{\delta}(x,y) \int_{S(x,y)}^{\infty} \overline{G}_1(\xi)\, \mathrm{d}\xi. \quad (9)
\end{aligned}
$$

The maximum worker surplus accumulates the usual annuity comprising the flow utility $(1 + r)\left[u(\overline{w}(x, y)) - c(x, y)\right]$ minus the unemployment annuity $rW_0(x)$. The extra term is the expected value of $\max\{\xi - S(x, y), 0\}$. If the mobility shock ξ is greater than the current surplus $S(x, y)$, the worker's option value is ξ and not the incumbent's reservation value $S(x, y)$.

We will also call $S(x, y)$ the match surplus. This is a slight misuse of language, as the term "match surplus" usually refers to the total production of a match minus the sum of what the various parties can produce on their own. This definition naturally arises in transferable utility models (for example, Lise et al., 2016). Here, there is no such definition of the surplus of a match that is independent of the way it is shared.

Utility is imperfectly transferable because firms' valuations are expressed in units of production (cash flow minus wage bill), whereas workers value wages through a utility function, so their job valuations are expressed in units of utility.

Firm profits and the value of a vacancy. Making use of Theorem 1 we can express firm profits (2) in terms of equilibrium wages:

$$\Pi_1(x, y, R) = \Pi_0(y) + \int_R^{S(x,y)} \frac{\mathrm{d}R'}{u'(w(x, y, R'))}. \tag{10}$$

Similarly, we can express the value of a vacancy (3) as a function of the equilibrium wage:

$$\begin{aligned} r\Pi_0(y) = {} & \frac{\lambda_0 L_0}{V} \int_{S(x,y)>0} \int_{-S(x,y)}^{0} \frac{\overline{G}_0(\xi)}{u'(w(x, y, -\xi))} \frac{\ell_0(x)}{L_0} \,\mathrm{d}\xi\,\mathrm{d}x \\ & + \frac{\lambda_1 L_1}{V} \iint_{S(x,y)>0} \int_{-S(x,y)}^{0} \frac{\overline{G}_1(\xi + S(x, y'))}{u'(w(x, y, -\xi))} \overline{\delta}(x, y) \frac{\ell_1(x, y')}{L_1} \,\mathrm{d}\xi\,\mathrm{d}x\,\mathrm{d}y'. \end{aligned} \tag{11}$$

In the next Section we will put these equilibrium properties to use for identification.

3 Identification

The model we have specified is particularly rich because it allows for the possibility of sorting in an environment with frictions, and it includes amenities and mobility shocks that can, in principle, allow for a complex structure of transitions and wage growth. But the value of such a rich specification hangs on our ability to identify the key components without relying on assumed parametric forms. This will define the empirical content of the model. We now discuss the identifiability of our model.

Throughout, we assume that the discount rate r and the flow utility function $u(w)$ are known. The deep parameters/functions of the model for which we need to show identification are the production function $f(x, y)$, the disutility (net of amenity) function $c(x, y)$, the utility of unemployment $b(x)$, the separation rate $\delta(x, y)$, the population measures of skill $\ell(x)$ and job productivity $n(y)$, the distribution of the mobility shocks for the unemployed $G_0(\cdot)$ and the employed $G_1(\cdot)$ and parameters λ_0, λ_1.[4]

[4]Note that the matching function $M(L, V)$ is only identifiable with aggregate (over time or space) variation of the unemployed and the vacancies. However, this function is only required for counterfactual simulations, and we can borrow it from the literature. Therefore, we focus on identifying λ_1 and λ_0.

To recover the structural parameters of the model, we assume the existence of matched employer-employee data over a finite number of periods T. Each individual i is associated with latent heterogeneity x_i and each employer j is associated with latent heterogeneity y_j, but both are unobserved by the econometrician. For each period t and individual i, we observe the employment status $e_{it} \in \{\mathrm{U}, \mathrm{E}\}$ and, if employed, the wage compensation w_{it} and the current employer j_{it}. We denote the joint density of these observables by $\mathbb{P}[j_{i1}, w_{i1}, e_{i1}, ..., j_{iT}, w_{iT}, e_{iT}]$, which can be obtained directly from the data. We also assume that a measure of the aggregate labor share is available.

We denote $m_{it} = \mathrm{EE}$ if the worker remains employed with the same firm between t and $t+1$, $m_{it} = \mathrm{EU}$ if the worker separates from a job into unemployment, and $m_{it} = \mathrm{JJ}$ if the worker changes employer. We denote $m_{it} = \mathrm{UU}$ if the worker is unemployed in both t and $t+1$ and $m_{it} = \mathrm{UE}$ if the worker moves from unemployment to employment.

3.1 Step 1: Distributions, transition probabilities, and wages

Assuming finite worker and firm latent types, we build on the results in Bonhomme, Lamadon, and Manresa (2019, BLM) on the identification of nonlinear Markovian wage equations and distributions with two-sided heterogeneity. To apply BLM's framework, we first establish the following result.

Lemma 1. *The sequential auction model of Section 1 produces a Markovian law of motion for wages and employment conditional on worker type, i.e.:*

$$\mathbb{P}[w_{t+1}, y_{t+1}, m_{t+1}|x, w_t, y_t, m_t, \Omega_{t-1}] = \mathbb{P}[w_{t+1}, y_{t+1}, m_{t+1}|x, w_t, y_t, m_t]$$

where $\Omega_t = \{w_\tau, y_\tau, m_\tau\}_{\tau=1,...,t}$ *is the information set in periods 1 to* t.

This step allows us to estimate a flexible reduced-form model of wages and mobility that nests the structural model. This will be one of the rare cases where indirect inference can be proven to be consistent. Usually, the identification of a structural model is based on intuitive considerations on how a moment should help identify a parameter. However, the injectivity of the binding function that links auxiliary parameters to structural parameters is rarely formally proven. Here, we start by proving that our model belongs to a large class of latent Markov models. In a second step, we will prove formally the identification of structural parameters from the auxiliary model. Very importantly, the identification of the distribution of latent worker and firm types will also be achieved in the first step. This is where usual moment-based estimators fail to produce convincing estimators of unobserved heterogeneity, as it is hard to separately identify different latent groups using aggregate moments. Hagedorn et al. (2017) were the first to propose an identification and estimation procedure that shares some of these ideas.

Assuming a finite number of worker types K, adapting BLM's method, we show that the cross-sectional distribution $\mathbb{P}[x, y, e_t = \mathrm{E}] = \ell_1(x, y)$, the moving probabilities $\mathbb{P}[m_t = m \,|\, x, y_t = y, e_t = \mathrm{E}] := p_m(x, y)$ for $m = \mathrm{EU}, \mathrm{EE}, \mathrm{JJ}$, the transition probability $\mathbb{P}[y_{t+1}, m_t = \mathrm{JJ} \,|\, x, y_t] := p_{\mathrm{JJ}}(y_{t+1}|x, y_t)$, the law of motion for within-job wages $\mathbb{P}[w_{t+1} \leq w \,|\, w_t, x, y_t, m_t = \mathrm{EE}] := F_{\mathrm{EE}}(w|x, y_t, w_t)$ and the distribution of wages after a move $\mathbb{P}[w_{t+1} \leq w \mid x, y_{t+1}, y_t, m_t = \mathrm{JJ}] := F_{\mathrm{JJ}}(w|x, y_t, y_{t+1})$ are non-parametrically identified from data on movers with two periods before and after the move.

We include unemployment as an extra state, allowing us to measure transition rates for each type of worker into and out of employment. We then recover the distribution among the unemployed $\mathbb{P}[x, e_t = \mathrm{U}] = \ell_0(x)$, the probability of exiting unemployment conditional on type $\mathbb{P}[m_t = \mathrm{UE} \,|\, x, e_t = \mathrm{U}] := p_{\mathrm{UE}}(x)$, the distribution of the destination firm conditional on type $\mathbb{P}[y_{t+1} \,|\, x, m_t = \mathrm{UE}]$, and the wage conditional on making that move $\mathbb{P}[w_{t+1} \leq w|x, y_{t+1}, m_t = \mathrm{UE}] := F_{\mathrm{UE}}(w|x, y_{t+1})$ (see the online appendix B.4 for complete details).

From this point on, we treat each of these distributions as known.

3.2 Step 2: Surpluses and wage functions

Let us denote by $R(x, y, w)$ the inverse of the wage function $R \mapsto w(x, y, R)$ for any given match type (x, y). We call it the worker surplus function. First, we prove the following identification result.

Lemma 2. *For each match (x, y), we identify its viability $\phi(x, y) = \mathbf{1}\{S(x, y) > 0\}$, and for each viable match (x, y), we identify the match surplus $S(x, y)$, the worker surplus function $R(x, y, w)$ and the equilibrium wage function $w(x, y, R)$ from the following observables:*

- *the match type distribution $\mathbb{P}\left[x, y_t = y, e_t = E\,\right] = \ell_1(x, y)$,*
- *the probability of within-job wage increase,*

$$\mathbb{P}\left[m_t = \mathrm{EE}, w_{t+1} > w|x, y_t = y, w_t = w, e_t = E\,\right] = \overline{F}_{\mathrm{EE}}(w|x, y, w),$$

- *the mobility probabilities* $\mathbb{P}[m_t = m | x, y_t = y, e_t = \mathrm{E}] = p_m(x, y)$, $m_t = \mathrm{EU}, \mathrm{JJ}$.

Viability $\phi(x, y)$ is directly identified from the fact that the matching sets are observed from the knowledge of $\ell_1(x, y)$ from the previous step: $\phi(x, y) = 1$ if and only if $\ell_1(x, y) > 0$.

The link between the distributions identified in Step 1 and the surplus functions is more subtle and is an important result. We deduce from equation (7) that differentiating the worker surplus $R(x, y, w)$ with respect to the wage w yields:

$$\frac{\partial R(x, y, w)}{\partial w} = \frac{(1 + r)u'(w)}{r + \delta(x, y) + \lambda_1 \overline{\delta}(x, y) \int \overline{G}_1 \left[R(x, y, w) - S(x, y')^+\right] \frac{v(y')}{V} \, \mathrm{d}y'}, \tag{12}$$

where the denominator is equal to r plus the probability of any change to the current state, which includes the probability of moving to unemployment, the probability of changing jobs, and the probability of staying but receiving a wage increase:

$$\begin{aligned} \delta(x, y) + \lambda_1 \overline{\delta}(x, y) \int \overline{G}_1 \left[R(x, y, w) - S(x, y')^+\right] \frac{v(y')}{V} \, \mathrm{d}y' \\ = \mathbb{P}[m_t = \mathrm{EU} | x, y_t = y, e_t = \mathrm{E}] + \mathbb{P}[m_t = \mathrm{JJ} | x, y_t = y, e_t = \mathrm{E}] \\ + \mathbb{P}[m_t = \mathrm{EE}, w_{t+1} > w | x, y_t = y, w_t = w, e_t = \mathrm{E}]. \end{aligned}$$

The probability of any of these changes conditional on a match (x, y) is known from Step 1 and therefore is identified.

The minimum wage in an (x, y) match is associated with a contract that yields zero surplus: $R(x, y, \underline{w}(x, y)) = 0$. We therefore obtain $R(x, y, w)$ by integrating (12) for wages w in the support $[\underline{w}(x, y), \overline{w}(x, y)]$:

$$R(x, y, w) = \int_0^w \frac{\partial R(x, y, w')}{\partial w} \, \mathrm{d}w'.$$

The match surplus follows as $S(x, y) = R(x, y, \overline{w}(x, y))$ (the maximum surplus gives the maximum wage). Lastly, $w(x, y, R)$ is identified as the inverse of $R(x, y, w)$.

It is important to note that equation (12) is quite general. For example, it will also be valid in a model without mobility shocks. The key property of the model that makes this possible is the fact that firms offer insurance when not matching outside offers, together with the fact that outside offers are independent of the current state. This implies that the rate at which the value increases is related to the probability that a change in state occurs.

3.3 Step 3: Mobility shock and vacancy distribution

Assumption 1. *We assume the following about the shock distributions:*

(a) G_1 *has zero median:* $G_1(0) = 1/2$*.*

(b) G_1 *belongs to a parametric family* $\mathcal{G}_1 = \{G_1(\xi;\theta)\}_\theta$*, where* θ *is identified from observations in any bounded interval.*

(c) The distribution G_0 *has support in* $(-\infty; 0]$*.*

(d) G_0 *belongs to a parametric family* $\mathcal{G}_0 = \{G_0(\xi;\theta)\}_\theta$*, where* θ *is identified from observations in any bounded interval.*

Assumption 1(a) imposes a normalization that allows separating λ_1 from $G_1(\xi)$. It only assumes that the distribution of the shock has a median of 0, i.e., it is centered. Assumption 1(c) is repeating the support restriction for $G_0(\xi)$ that rules out positive preference shocks when coming out of unemployment. The assumptions 1(b) and 1(d) are restrictions on the families of distributions that we can consider. This is an extrapolation assumption. It covers a large set of functions, for example, the family could include both Normal and Logistic, or it could include any polynomial transformation of ξ inside a Logit.

Lemma 3. *Define* $\overline{S} = \max_{x,y} S(x,y)$*,*

1. *Under Assumption 1(a),* $G_1(\xi)$ *is nonparametrically identified on* $\xi \in [-\overline{S}, 0]$ *from* $F_{\text{JJ}}(w|x,y,y')$ *and knowledge of* $R(x,y,w)$ *and* $S(x,y)$*.*
2. *In addition, with Assumption 1(b),* $G_1(\xi)$ *is identified everywhere.*

This lemma establishes that we can learn very flexibly about $G_1(\xi)$. The result comes from the following characterization of job-to-job wages. Conditional on a move from y to y', the only uncertainty is about the mobility shock ξ. We know that the new match surplus $S' \equiv S(x,y') > 0$ and that $S' + \xi > S \equiv S(x,y)$. The new contract has value $R' = \max\{S - \xi, 0\}$ and the workers also collect the mobility shock. The new wage is given by $w(x,y',R')$.

Hence, for the minimum transition wage $w = w(x,y',0) := \underline{w}(x,y')$,

$$F_{\text{JJ}}(w|x,y,y') = \frac{\overline{G}_1\left[S(x,y)\right]}{\overline{G}_1\left[S(x,y) - S(x,y')\right]},$$

and, for any positive wage $w \in \left]\underline{w}(x,y), \overline{w}(x,y)\right[= w(x,y',R')$ with $0 < R' = R(x,y',w) < S'$,

$$F_{\text{JJ}}(w|x,y,y') = \frac{\overline{G}_1\left[S(x,y) - R(x,y',w)\right]}{\overline{G}_1\left[S(x,y) - S(x,y')\right]}.$$

Intuitively, for realized matches, the shock moves the compensation that a poacher has to offer the new hire. The distribution of starting wages conditional on a move traces out the distribution of the preference shock. We have already identified $F_{\mathrm{JJ}}(w|x,y,y')$ in Step 1 and $R(x,y,w) \in [0, S(x,y)]$ from Step 2. We vary $S(x,y) - R(x,y',w)$ and $S(x,y) - S(x,y')$, and make use of the centering $G_1(0) = 1/2$. When $y = y'$, we identify $\overline{G}_1 [S(x,y) - R(x,y,w)]$ continuously for all $R(x,y,w)$ between 0 and $S(x,y)$. When $S(x,y) < S(x,y')$ we identify $\overline{G}_1 [S(x,y) - S(x,y')]$ when $R(x,y',w) = S(x,y)$. Identification of G_1 below 0 is based on assumptions on the support of firm types. If $n(y)$ is a discrete distribution, then the identification of G_1 below 0 occurs only at the points $S - S'$.

Lemma 4. *The probability of separation $\delta(x,y)$, the rate of meeting on the job λ_1 and the density of vacancies $v(y')/V$ are identified from $p_{\mathrm{JJ}}(y'|x,y)$, $p_{\mathrm{EU}}(x,y)$ and $G_1(\xi)$.*

This result comes from the expression for the transition $p_{\mathrm{JJ}}(y'|x,y)$, that relates the probability of moving to the difference in surpluses in the distribution of the preference shock given by:

$$p_{\mathrm{JJ}}(y'|x,y) = \bar{\delta}(x,y)\lambda_1 \frac{v(y')}{V} \overline{G}_1(S(x,y) - S(x,y')). \tag{13}$$

Since this is observed for all (x,y,y') over viable matches and $\overline{G}_1(\xi)$ and $S(x,y)$ are known, we can use relative flows $p_{\mathrm{JJ}}(y'_1|x,y)/p_{\mathrm{JJ}}(y'_2|x,y)$ to recover $v(y')/V$.

We note here that equation (13) allows us to identify G_1 on a larger support than Lemma 4 since we can use transitions where $S(x,y) - S(x,y') > 0$. This allows for relaxation of the use of Assumption 1(b). Combining this with $p_{\mathrm{EU}}(x,y)$, the layoff rate $\delta(x,y)$ is identified by the conditional probability of separation implied by the model.

Lemma 5. *Given $p_{\mathrm{UE}}(x,y)$, $F_{\mathrm{UE}}(w,x,y)$ and $v(y)/V$,*

1. *$\lambda_0 \overline{G}_0(\xi)$ is non parametrically identified on $[-\bar{S}, 0]$,*
2. *With Assumption 1(d) it is identified everywhere.*

This result comes from the expression of the flow out of unemployment together with the wage conditional on coming out of unemployment. Importantly, we can learn $G_0(\xi)$ very flexibly and, in particular, for any counterfactual where the surpluses do not move outside of $\bar{S}$, we do not need the extrapolation Assumption 1(d).

3.4 Step 4: Production function and amenities.

Lemma 6. *Under Assumption 1, and knowledge of $\lambda_1 G_1(\xi)$, $\lambda_0 G_0(\xi)$, $S(x,y)$ and $v(y)/V$ then $f(x,y)$, $\tilde{c}(x,y) = c(x,y)+b(x)$, V and $n(y)$ are identified from knowledge of the aggregate labor share.*

Establishing the identification of $f(x,y)$ leverages important properties of the model. First, we can express the output of the match using the equilibrium wage $w(x,y,R)$, the surplus $S(x,y)$, and the value of a vacancy $\Pi_0(y)$:

$$f(x,y) = w(x,y,S(x,y)) + \frac{r}{1+r}\Pi_0(y), \tag{14}$$

where the only unknown part is $\Pi_0(y)$. Using equation (11) we see that $\Pi_0(y)$ is only unknown up to the total number of vacancies V.[5] Hence, the match output is known up to one coefficient. Integrating the output of the match against the already identified distribution of matches $\ell_1(x,y)$ expresses V as a function of the labor share and known quantities:

$$\text{labor share} = \frac{\mathbb{E}\left[w_{it}\right]}{\mathbb{E}\left[w(x_i, y_{j(i,t)}, S(x_i, y_{j(i,t)})) + \frac{1}{V}\frac{rV}{1+r}\Pi_0(y)\right]}. \tag{15}$$

This is a substantive result, since we did not assume transferable utility where knowledge of $S(x,y)$ would have directly provided us with $\Pi_0(y)$ up to scale. Here, instead, we used the property of the optimal contract that relates the Pareto frontier to the marginal utility of the wage, as captured in the equation (11).

The final part of Lemma 6 shows the identification of the disutility term $\tilde{c}(x,y) = c(x,y) + b(x)$. Without external information on the value of unemployment, we can only treat unemployment as the external good and measure each job type relative to this option. This means that we identify $\tilde{c}(x,y) = c(x,y) + b(x)$, which captures the disutility of a particular job, net of any amenities, and net of the forgone leisure and home production. We can identify this from the surplus equation (9) where everything is known other than the term $\tilde{c}(x,y)$. Thus, $\tilde{c}(x,y)$ is identified as the component of the maximum flow surplus that is not already accounted for by the maximum wage and the expected value of future mobility shocks.

[5] Equation (11) equates the value of a vacancy with the expected gain from matching. If an additional period of vacancy incurred a type specific cost $k(y)$, this vacancy cost would not be identifiable.

Taking stock. In this section we have laid out in a transparent way the identification of the model structure. The surplus of any (x, y) match is identified by aggregating the marginal utility of the wages observed in the match, weighted using the expected duration in the state. The expected duration encodes the revealed preference information of the workers. The production function is identified by the maximum wage paid in an (x, y) match plus the opportunity cost of a vacancy for a firm. The disutility of labor net of amenities term is identified by the discrepancy between the surplus and the maximum wage in an (x, y) match. In other words, $\tilde{c}(x, y)$ is needed to rationalize any systematic differences in the firm rankings of workers based on mobility-weighted wages compared to wage-based ranking alone. Finally, the distributions of mobility shocks are identified by idiosyncratic worker mobility that is inconsistent with the systematic rankings implied by the surplus, along with the distribution of wages at employment transitions.

4 Data, Estimation and Results

Before we describe the data, we give an overview of our estimation procedure. While we build on the identification argument, estimation requires us to make some functional-form assumption. In particular, we will assume that $G_1(\xi)$ is a logistic and $G_0(\xi)$ is a logistic truncated above by zero. For Step 1, we specify a flexible discrete heterogeneity model directly following BLM. Next, we make direct use of the logistic specification to estimate surpluses and mobility parameters up to knowledge of the parameters of $G_1(\xi)$ and $G_0(\xi)$ from the mover probabilities using Equation (13). We finally use these scaled surpluses together with the wage moments of F_{EE}, F_{UE} and F_{JJ} to estimate the remaining parameters. Importantly, the production function and disutility remain completely unrestricted.

Data We use matched employer-employee data from Sweden. The data comprise annual tax records for the universe of jobs in Sweden. Each record provides information on the start and end month of the spell in each year, an employer identifier, an employee identifier, and the total compensation for the year.

Using the monthly spell level information, we construct transition rates at the quarterly frequency and associated monthly earning equivalent based on the number of months worked and total compensation. We track workers in and out of recorded unemployment and derive their employment state accordingly. We use five years of data from 2000 to 2004, and included all workers under the age of 50. See the online appendix B.5 for additional details.

Estimation of the reduced-form model We start by classifying the firms following BLM. We group firms based on the empirical CDF of wages in the cross section. Given this classification, we estimate the distribution of types and wages, as well as the probability of moving using maximum likelihood.

We specify flexible probability models for transitions out of unemployment, wages out of unemployment, wages while on the job, the probability of moving to a new firm, and the wage conditional on a move. As in the rest of the paper, we use x to denote worker types and y to denote firm types and we focus on the case with discrete types.

We continue to denote $\ell_0(x)$ as the probability that a worker is unemployed and of type x, and $\ell_1(x, y)$ as the probability that a worker is of type x and matched with a firm of type y. We leave these probabilities completely unrestricted. We also leave unrestricted the probability of leaving unemployment, the probability for a worker of type x to join a firm of type y, which we denote $p_{\mathrm{UE}}(x, y)$. Finally, we specify the hiring wage for a worker of type x who comes out of unemployment and joins the firm y as a log-normal with unrestricted mean $\mu_{\mathrm{UE}}(x, y)$ and variance $\sigma_{\mathrm{UE}}(x, y)$. Hence the probability of observing a wage w for a worker joining firm y in the data is given by:

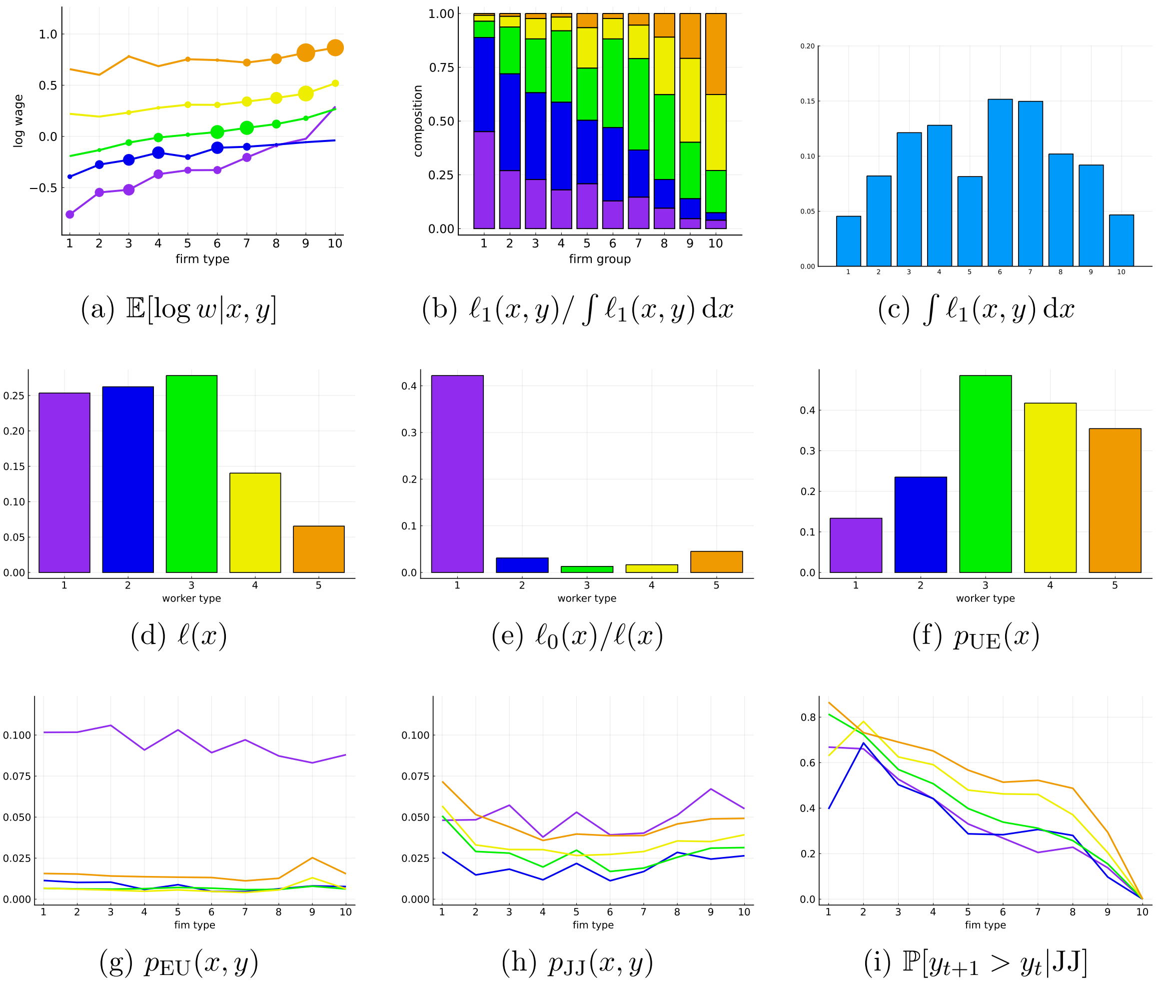
$$\mathbb{P}[y_t{=}y, w_t{=}w \mid x, m_{t-1}{=}\mathrm{UE}] = p_{\mathrm{UE}}(x, y)\mathcal{N}\big(w, \mu_{\mathrm{UE}}(x, y), \sigma_{\mathrm{UE}}(x, y)\big)$$

where $\mathcal{N}$ is the density of the log-normal distribution. The evolution of wages on the job is specified as an auto-regressive process with (x, y) specific intercept and a common auto-regressive coefficient:

$$\mathbb{P}[w_t \mid x, y, w_{t-1}, m_{t-1}{=}\mathrm{EE}] = \mathcal{N}\big(w_t - \gamma w_{t-1}, \mu_{\mathrm{EE}}(x, y), \sigma_{\mathrm{EE}}(x, y)\big).$$

The probability of a transition from job to job is left unrestricted as a function of x, y as is the type of destination firm y', and is denoted as $p_{\mathrm{JJ}}(x, y, y')$. Importantly, following the model presented in the previous section, the wage conditional on move is a distribution which depends only on the worker and firm types, but conditional on these types is independent of the wage before the move. We then use a log-normal distribution and let its mean and variance be unrestricted functions of x, y, y'.

We jointly estimate the allocations $\ell_0(x)$ and $\ell_1(x, y)$, the mobility parameters



(a) $\mathbb{E}[\log w|x, y]$ (b) $\ell_1(x, y)/\int \ell_1(x, y)\,\mathrm{d}x$ (c) $\int \ell_1(x, y)\,\mathrm{d}x$

(d) $\ell(x)$ (e) $\ell_0(x)/\ell(x)$ (f) $p_{\text{UE}}(x)$

(g) $p_{\text{EU}}(x, y)$ (h) $p_{\text{JJ}}(x, y)$ (i) $\mathbb{P}[y_{t+1} > y_t|\text{JJ}]$

Figure 1: Type distributions and conditional transitions (Step 1 estimates)

Notes: Firm and worker types are ordered by mean wages.

$p_{\text{UE}}(x, y)$, $p_{\text{JJ}}(x, y, y')$, and $\delta(x, y)$, and the parameters of the wage equation $\mu_{\text{UE}}(x, y)$, $\sigma_{\text{UE}}(x, y)$, $\mu_{\text{EE}}(x, y)$, $\sigma_{\text{EE}}(x, y)$, $\mu_{\text{JJ}}(x, y, y')$, $\sigma_{\text{JJ}}(x, y, y')$, γ using maximum likelihood. Given n_x worker types and n_y firm types, this amounts to a number of parameters of the order of $3n_x \cdot n_y^2$. In estimation, we focus on $n_x = 5$ and $n_y = 10$.

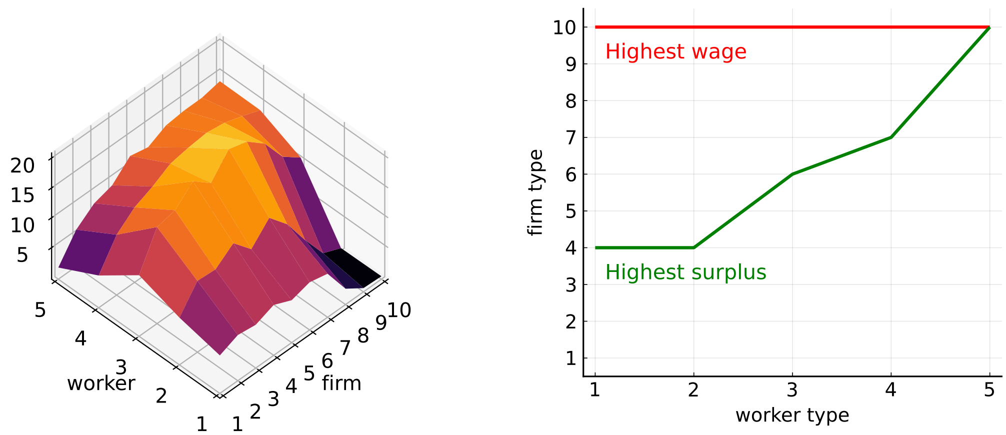
Reduced Form Results From the Step 1 estimates, we can already see clear differences between the outcomes of different types of workers, indicating a pattern of sorting of workers between firms. In Figure 1 we show most of the estimated objects from Step 1. In panel (a) we plot, for each worker type x, the mean log wage when matched to a firm type y. For the purposes of the figures, we order the worker types by their mean wage and the firm types by the mean wage they pay. This ordering appears sensible and produces a clear ordering of firms where all workers are more or less paid more as we move to higher firm types. The figure also reveals a relatively stable ordering of workers, with the exception of the lowest worker type by the top two firm types, where the data are sparse.

In panel (b), we plot the composition of worker types employed in each type of firm, and in panel (c), we plot the share of total employment of each type of firm in the cross section. The cross-sectional distribution shows a clear and strong pattern of separating high-paid workers into high-paying firms. In panels (d), (e), and (f), we plot the estimated proportions of workers by type, the type-specific unemployment rate, and the type-specific job-finding rate from unemployment. In panels (g), (h), and (i), we plot the probability of moving from employment to unemployment, the probability of changing job, and the probability of moving to a higher firm type conditional on changing job.

The estimates in Step 1 produce an interpretable picture of what distinguishes the worker types from each other. Take, for example, the lowest-type workers, who make up about a quarter of the population. Not only do they have the lowest average earnings when employed, they also have the highest unemployment rate (over 40%), which is generated by the lowest probability of moving from unemployment to employment, combined with a probability of separating into unemployment from any type of firm that is about 10 times higher than the average of the other types of workers. These workers also have high job-to-job transition probabilities, but have the lowest probability of moving to a higher paying firm conditional on changing jobs. On the other hand, workers with the highest type, about 7% of the population, have the highest wages when employed, unemployment rates below 5%, a high probability of moving from unemployment to employment and a low probability of separating from employment to unemployment. They also have a high job-to-job transition rate, but when they change jobs, they have the highest probability of moving to a higher paying firm.

To summarize, it is clear from the Step 1 estimates that worker types differ across important wage and mobility dimensions, and that these differences generate a distribution of worker types across firm types that is strongly positively sorted.[6]

[6] In Online Appendix Figure 6 we plot the stationary distribution of worker types across firm



(a) Surplus $S(x, y)$ (b) Firm ranking

Figure 2: Surplus and Firm Rank (Step 2)

Notes: We plot mean the surplus $S(x, y)$ and the comparison of firm rankings based on wages and surplus. Estimates are from step 2.

Estimation of surpluses, vacancy rates, and meeting rates. We choose simple parametric families for G_1 and G_0:

$$G_1(\xi) = \frac{1}{1 + e^{-\rho_1 \xi}}, \quad G_0(\xi) = \frac{2}{1 + e^{-\rho_0 \min\{0,\xi\}}}.$$

We assume that the preference shocks are distributed according to a logistic distribution for G_1 and a logistic truncated above at 0 for G_0. We allow each to have its own scale parameter, ρ_1 and ρ_0. For each (ρ_0, ρ_1), we estimate surpluses, vacancies $v(y)/V$ and λ_0 and λ_1 to maximize the likelihood of the transitions estimated in the reduced form model (see Online Appendix B.6 for the likelihood). We constrain the surplus to be positive whenever $\ell_1(x, y) > 0$ as implied by the theory. Given such surplus, we compute the mean full-year employment wage growth of stayers, as well as the difference in full-time employment wages after a move from unemployment versus from another firm, as captured in the following moments:

$$m_1 = \mathbb{E}\left[\log w_t - \log w_{t-4} \middle|\, s_t{=}s_{t-4}{=}1,\ m_{t-4}{=}\mathrm{EE}\right],$$

$$m_2 = \mathbb{E}\left[\mathbb{E}\left[\log w_t | x, y, s_t{=}1, m_{t-4} = \mathrm{JJ}\right] - \mathbb{E}\left[\log w_t | x, y, s_t{=}1, m_{t-4} = \mathrm{UE}\right]\right],$$

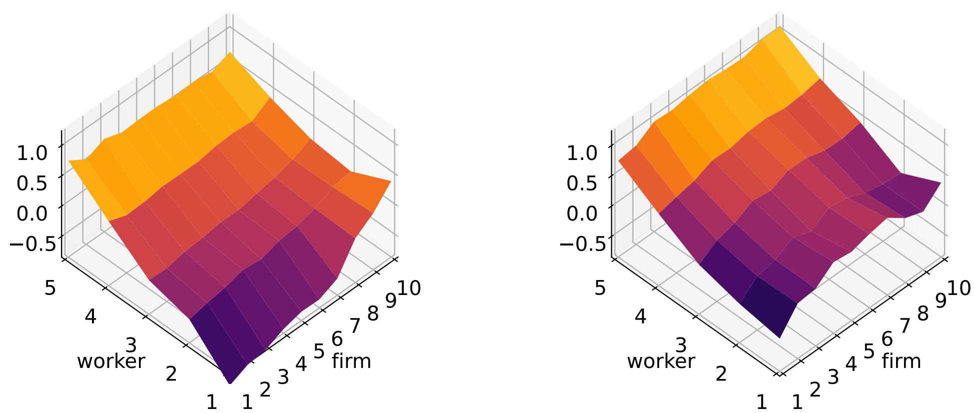
types implied by the model, along with the cross-sectional distribution in the data estimated in Step 1. The distributions align remarkably well, which supports the restriction of attention to the steady state of the model.

where t is quarter and s_t=1[m_{t-1}=m_{t-2}=m_{t-3}=EE] denotes a full time employment year. In the data $\hat{m}_1 = 0.01$ and $\hat{m}_2 = 0.083$. The procedure matches them exactly. We obtain an estimate of $\hat{\rho}_1 = 0.42$ and $\hat{\rho}_0 = 0.07$. This suggests the presence of a larger variance in cost when leaving unemployment than when moving from another firm.

In Figure 2(a) we plot the estimated surplus $S(x, y)$. We note several interesting patterns in the surplus implied by transitions between firm types. First, while mean wages suggest an apparent common ranking of firms by workers, mobility patterns do not. In Figure 2(b) we plot the highest paying firm and the highest surplus firm by worker type. Although there is unanimous agreement among worker types about which firms are high-pay, there is complete disagreement by worker type as to which firms provide high surplus. These results accord well with the empirical findings of Lentz et al. (2023) that inferences about sorting using only information on wages are very different from inferences that use information on mobility patterns. In the next sections, we turn to estimating the underlying structure that provides a fully coherent interpretation of these results.

Disutility of labor and the production function. We use equation (7) and identified quantities to obtain an estimate of the type-specific disutility of labor $\tilde{c}(x, y)$. We then have all the elements to construct $\Pi_0(y)$ using the equation (11) up to one scale value, which is the total number of vacancies. We obtain this scale value by matching a labor share of 0.75 using equation (15). We then use equation (14) to recover the production function $f(x, y)$.

We plot the estimates of $\tilde{c}(x, y)$ and the log of $f(x, y)$ in Figure 3. There are several patterns of note here. First, the average disutility of work is monotonically increasing in worker type; higher worker types have a higher opportunity cost of work. For the lowest worker types, the disutility of labor is increasing in the firm type, indicating that while these firms pay well, they are unattractive for low-type workers on their non-wage dimensions. This is in contrast to the highest worker type, where the disutility of labor is high, but effectively independent of the firm type. Finally, the estimates reveal striking monotonicity in both dimensions x and y (with the exception of the lowest worker type paired with the highest firm, where there is little data). Turning to the estimates of the production function, we note the striking monotonicity in both x and y, again with the exception of the pairing



(a) Disutility of work $\tilde{c}(x, y)$ (b) Log production $\log f(x, y)$

Figure 3: Disutility net of amenities and production function (Steps 4 and 5)

of the lowest worker type and the highest firm type, where very few matches occur. The monotonicity present in the estimates of $\tilde{c}(x, y)$ and $f(x, y)$ supports our choice of ordering workers and firms according to the mean wages in Figure 1(a).

5 Wage and value decomposition

In this section, we use the estimated model to decompose the variance of log wages, we assess the quantitative importance of mobility shocks and interpret well-documented empirical regularities through the structure of the model.

Wage variance decomposition and compensating differentials. We begin our decomposition of the variance of log wages with the standard within and between worker decomposition:

$$\underbrace{Var(\log w)}_{\text{total}} = \underbrace{\mathbb{E}[Var(\log w|x)]}_{\text{within worker: 31\%}} + \underbrace{Var(\mathbb{E}[\log w|x])}_{\text{between worker: 69\%}}. \tag{16}$$

Turning first to the within worker (within x-type) term, we decompose this into interpretable sources. For a given worker, the wage variation is generated from three different sources. The first is the variation in wages due to compensating differentials. In the face of non-wage differences between jobs, different firms must pay the same worker different wages to deliver the same value (Rosen, 1986). Labor market frictions allow for two additional sources of variation. The presence of search frictions allows for the coexistence of firms that offer different values to the same worker (Mortensen, 2003). The effect of search frictions is amplified by sequential auctions, since this allows a firm to offer the same worker a different value over time as their outside option evolves. This is the variation within a worker-firm match that is induced by the sequential auction (Postel-Vinay and Robin, 2002). We will refer to these three sources of within worker-type variation as compensating differentials, search frictions, and sequential auctions; or Rosen, Mortensen, and PVR sources for short.[7]

For a fixed worker type x, we decompose wages first into the variation within and between a fixed surplus value R, and second, we decompose the variation between R into the variation between and within firms. To simplify notation we first define $\mu_R = \mathbb{E}[\log w|R, x]$ and write:

$$\underbrace{Var(\log w|x)}_{\text{within worker}} = \underbrace{\mathbb{E}[Var(\log w|R, x)|x]}_{\text{within } R} + \underbrace{Var(\mu_R|x)}_{\text{between R}} \tag{17}$$

$$= \underbrace{\mathbb{E}[Var(\mathbb{E}[\log w|R, y, x]|, x)|x]}_{\text{(a) Rosen}} + \underbrace{Var(\mathbb{E}(\mu_R|y, x)|x)}_{\text{(b) Mortensen}} + \underbrace{\mathbb{E}[Var\,[\mu_R|y, x]\,|x]}_{\text{(c) PVR}}.$$

The first term captures that fact that different firms pay different wages to deliver the same present value to a given worker; the Rosen compensating differential contribution. The second term captures the variance between employers in average wage net of the compensating differential. In other words, this captures the fact that some firms can offer larger values than others (including differences due to worker preferences, match separation rate, and the productivity of the firm); the Mortensen search friction contribution. The final term captures the wage variation for a given worker, within a given firm, net of the compensating differential; the PVR sequential auction contribution.

We present this decomposition in Figure 4 for each worker type separately, as well as for the average over the worker types. Each bar presents the share of the wage variance within the worker that is attributed to (a) Rosen compensating differentials, (b) Mortensen frictions, and (c) sequential auctions of PVR. Looking first at the average, the share of wage variation for a typical worker due to compensating differentials is slightly below 50%. This is an important finding as it implies that for the typical worker, half of the variation in wages we observe can be attributed to compensating differentials that do not reflect utility differences. Second, there are

[7] Our labeling generalizes the one adopted by Sorkin (2018) to include the within match dispersion resulting from the sequential auction between incumbent and poaching firms.

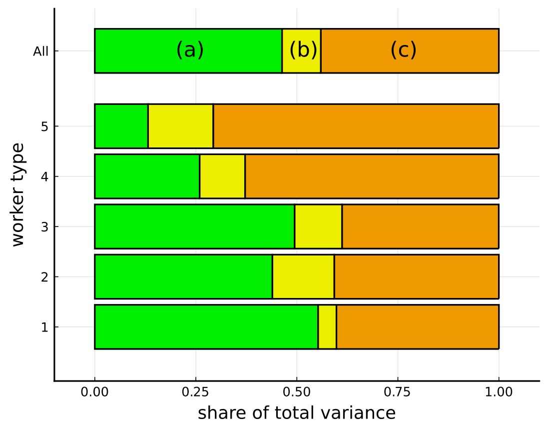


Figure 4: Within worker wage variance contribution

Notes: We plot the three terms of the equation (17) for each type of worker and for the average worker in the cross section. The bars represent the share of the variance within the worker accounted for by (a) Rosen compensating differential, (b) Mortensen frictions, and (c) sequential auctions of PVR. These are as a share of the within-worker variance, which amounts to 24% for the total variance.

substantial differences between workers. At the bottom of the distribution, there is substantially less variation in value compared to the variation in wages. About 55% of the wage variation for the lowest group is the compensating differential and occurs at a constant utility value. At the top, only 10% of the wage variation is due to compensating differentials. High-paid workers face important utility differences between employers and important within job wage dynamics resulting from the offer-counter offer mechanism.

We next decompose the between-worker variance into the within- and the between-firm variance:

$$\underbrace{Var(\mathbb{E}[\log w|x])}_{\text{between worker: 69\%}} = \underbrace{\mathbb{E}[Var(\mu_x|y)]}_{\text{within firm, between worker: 51\%}} + \underbrace{Var(\mathbb{E}[\mu_x|y])}_{\text{between firm, sorting 18\%}} .$$

The first term, $\mathbb{E}[Var(\mu_x|y)]$, captures the average within-firm variance of average worker values. This reflects the fact that even in the case where all firms are identical and pay identical wages, firms will hire a distribution of workers, and each worker has a different average market wage. The second term, $Var(\mathbb{E}[\mu_x|y])$, is the between firm variance in wages net of firm-specific pay policies. This is the contribution to the total variance of wages associated with the fact that different firms hire workers with different average market values. This speaks to the fact that different firms have different worker compositions and directly reflects the sorting of workers across firms.

In summary, most of the wage variance comes from differences between workers, which represents 51% of the overall wage variation. The sorting of workers to firms accounts for another 18% of the overall variance. The remaining variance occurs within the worker type and can be split as 14% due to compensating differentials at a fixed utility level, 3% resulting from different firms offering different values to the same workers, and the remaining 14% resulting from sequential auctions. The last two are the combined effect of market frictions, indicating that search frictions account for 17% of wage variation for a typical worker, although this share varies substantially between worker types.[8]

The quantitative role of the preference shock An important part of our framework comes from the introduction of the preference shock. To quantify its role, we conducted two simple exercises. First, using the model, we find that for 81% of the meetings between employed workers and poaching firms, the mobility result would be the same if the preference shock were zero. Second, to evaluate the importance of the preference shocks for wage variation, we simulate the model, replacing each preference shock with zero (the mean for outside offers). This simulation produces wages that are solely determined by worker type, firm type, and poaching firm type, rather than by a combination of types and preference shock. In this counterfactual, focusing on employed workers who have received at least one offer, the wage variance within (x, y) matches amounts to 86% of the baseline wage variance. In summary, the model attributes 19% of job-to-job moves and 14% of wage variation within matches to the preference shock.

The consequences of mobility Next, we use the structure of the model to compute the average gain and loss associated with moves by worker type. We start by reporting the change in the average wage of the employer in the first row of Table 1. On average, worker types 1, 2, and 3 move to firms with lower average wages, while worker types 4 and 5 move to firms with higher average wages. This confirms what

[8]An alternative to our decomposition would be to report the measures of worker/firm heterogeneity and sorting that come from a two way best linear predictor decomposition in the spirit of AKM. Applying AKM to model simulated data, we find that the shares of wage dispersion are worker fixed effect 53%, firm fixed effect 11%, covariance 20.5%, nonlinearities 1.7%, and residual of 14.1%. While not directly comparable to our decomposition, the shares attributable to worker heterogeneity and sorting (covariance) are strikingly similar.

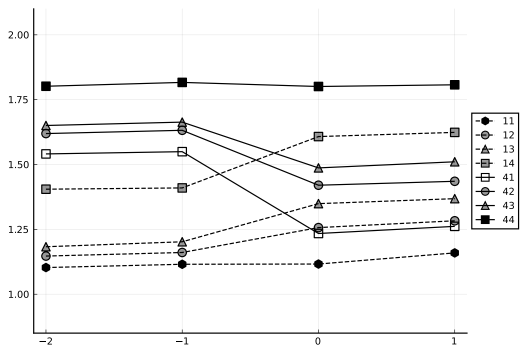
Table 1: Gains and losses upon moves

worker type	1	2	3	4	5	all
$\Delta\mathbb{E}[\log w \vert y]$	-0.2%	-1.4%	-2.7%	2.6%	7.4%	-0.1%
$\Delta\mathbb{E}[\log w \vert x, y]$	-0.3%	0.2%	-0.6%	0.8%	1.5%	0.1%
$\Delta \log f(x, y)$	1.2%	2.5%	2.6%	1.3%	1.8%	2.0%
$\Delta\tilde{c}(x, y)$	-1.6%	-1.1%	-1.8%	0.1%	-0.6%	-1.1%

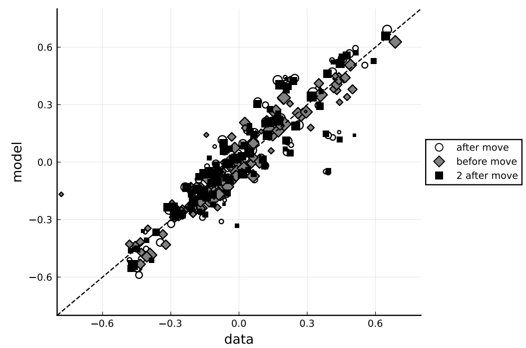
we reported in Figure 1(i) that low-type workers tend to move down the hierarchy of firm type according to the average wage. This pattern continues to hold when we condition on the match type. In the second row, we see that low-type workers tend to move to firms that pay them less, and high-type workers tend to move to firms that pay them more. Using the model, we can also look directly at the implied changes in output and amenities associated with a job change. The patterns are quite different here. On average, all types of workers move to jobs with higher productivity and better amenities (lower disutility). Overall, the mobility from job to job generates a 2% gain in output. Similarly, moves tend to improve amenities, with an average gain of 1.1%. Across types, we notice that the lowest-type workers appear to be improving amenities relatively more than output after a move.

Empirical regularities through the lens of the model. We use the model to provide a structural interpretation of several empirical regularities that have been documented. We start with the shape of the job-mover event study from Card et al. (2013). In Figure 5 we reproduce this study based on our estimated model.

We notice several important features. First, as typically documented, the movers from the top to the bottom quartile experience an earnings loss that is not alleviated in period 1. Sequential contracting provides a rationale for why workers might accept a wage cut upon a move, expecting that they would later recover through outside offers. In our model, however, compensating differentials arising from the $c(x, y)$ function can generate moves to higher value firms that actually pay permanently lower wages. As seen in Figure 4, the compensating differentials account for approximately half of the overall variance in the wage between workers. Second, the event study exhibits the usual symmetry in wage changes between upward and downward moves. We note



(a) Movers event study



(b) Fit around the move

Figure 5: Event Study plot.

Notes: Figure a) reproduces the event study plot from Card et al. (2013), it plots yearly earnings before and after a move for different firms, grouped by quartiles of earnings. Figure b) shows the fit of the model for annual earnings before and after a move for all pairs of firm types.

that this is also consistent with the important role of compensating differentials.

A second empirical regularity is the fact that wages appear more or less log-additive as presented in Figure 1(a) yet the surplus function shows that workers rank firms very differently and Figure 1(b) shows a lot of sorting in equilibrium. Without compensating differentials and our $c(x, y)$, these patterns would be difficult to reconcile. However, the surplus represents a combination of wage and amenity. This means that the model can rationalize both the wages and the surplus by having both $f(x, y)$ and $c(x, y)$. The combination of log-additive wages and strong sorting leads to estimated dis-amenities of work that increase across firm types for the low type workers to generate positive sorting.

A third empirical finding is that the firm in which a worker was employed before a move does not have a strong effect on the wage after the move (Di Addario, Kline, Saggio, and Sølvsten, 2023). This suggests that wages appear to be generated more by wage posting than by sequential contracting. An important aspect in a model like ours, where sorting is modeled structurally, is that the shape of the surplus will lead to heterogeneous effects of the previous firms on different workers. This will tend to muddle the average effect of the previous firm. Indeed, a linear additive formulation on our simulated data finds little role for the previous employer. However, structurally, there is indeed an effect.

6 Conclusion

In this paper, we develop an equilibrium model of the labor market with heterogeneous workers and firm types, type-specific production and disutilities/amenities, mobility preference shocks, and optimal contracting. We show that the model is non-parametrically identified from a short panel of matched employer-employee data. We find that for the lowest type workers, compensating differentials account for about half of the within-worker wage variation, but less than 10% for the highest type of worker. Finally, we use the structural model to interpret several empirical regularities that, in the absence of the model, are difficult to reconcile.

The model is very amenable to conducting counterfactual policy experiments and assessing the efficiency and distributional effects of, for example, changing the progressivity of taxes, earned income tax credits, or changing minimum wages. We estimate that compensating differentials for amenities play an important role in explaining wage variation. Policy changes that affect the relative value of wages and amenities will directly affect how workers value different jobs and will affect mobility patterns.

References

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

Arcidiacono, P., A. Gyetvai, A. Maurel, and E. Jardim (2023): "Identification and Estimation of Continuous-Time Job Search Models with Preference Shocks," Unpublished manuscript.

Balke, N. and T. Lamadon (2022): "Productivity Shocks, Long-Term Contracts, and Earnings Dynamics," *American Economic Review*, 112, 2139–77.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846.

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

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

Burdett, K. and D. T. Mortensen (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 257–273.

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

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

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

Friedrich, B., L. Laun, and C. Meghir (2022): "Earnings dynamics of immigrants and natives in Sweden 1985–2016," *Quantitative Economics*, 13, 1803–1847.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying equilibrium models of labor market sorting," *Econometrica*, 85, 29–65.

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 112, 169–212.

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

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

Morchio, I. and C. Moser (2023): "The Gender Pay Gap: Micro Sources and Macro Consequences," NBER Working Paper No. 32408.

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

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70, 2295–2350.

Rosen, S. (1986): "Chapter 12: The theory of equalizing differences," Elsevier, vol. 1 of *Handbook of Labor Economics*, 641–692.

Sattinger, M. (1993): "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 31, 831–880.

Shimer, R. and L. Smith (2000): “Assortative Matching and Search,” *Econometrica*, 68, 343–369.

Sorkin, I. (2018): “Ranking firms using revealed preference,” *The Quarterly Journal of Economics*, 133, 1331–1393.

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

A Appendix

A.1 Proof of Theorem 1: The optimal contract

In all the following, the outside options W_0, Π_0 are given. We will consider worker values on a compact set, as a firm should never offer more than the present value of output. We define $\mathbb{S} = [0, \max_{xy} f(x,y)/r]$. We also define the control spaces, that is, $R_1 \in L_\infty(\mathbb{S} \times \mathbb{R})$ and $B \in L_\infty(\mathbb{R})$. $R_0 \in \mathbb{S}$ and $w \in [\underline{w}, \bar{w}]$, supposed large enough not to bind in firm choices. Finally, we give the firm the option to fire the worker by choosing $\tilde{\delta} \in [\delta, 1]$, which is only useful away from the optimal solution (which will always choose $\tilde{\delta} = \delta$). We can define the following operator T that maps $\Pi_1(R)$ to $\hat{\Pi}_1(R) = T[\Pi_1](R)$ as:

$$\hat{\Pi}_1(R) = \sup_{\{w,R_0,R_1,B,\tilde{\delta}\}} \Big\{ f - w + \frac{\tilde{\delta}}{1+r}\Pi_0 + \frac{(1-\tilde{\delta})(1-\lambda_1)}{1+r}\Pi_1(R_0) + \tag{18}$$
$$\frac{(1-\tilde{\delta})\lambda_1}{1+r} \iint \Big[\mathbf{1}_{\{B(\xi) \geq B'^{+}+\xi\}} \Pi_1(R_1(B',\xi)) + \mathbf{1}_{\{B(\xi) < B'^{+}+\xi\}} \Pi_0 \Big] \, \mathrm{d}F(B'|\xi) \, \mathrm{d}G_1(\xi) \Big\},$$

subject to the promise keeping constraint,

$$R = u(w) - c - \frac{r}{1+r} W_0 + \frac{(1-\lambda_1)(1-\tilde{\delta})}{1+r} R_0 + \frac{(1-\tilde{\delta})\lambda_1}{1+r} \iint \Big[\mathbf{1}_{\{B(\xi) \geq B'^{+}+\xi\}} R_1(B',\xi)$$
$$+ \mathbf{1}_{\{B(\xi) < B'^{+}+\xi\}} \max\{B(\xi), \xi\} \Big] \, \mathrm{d}F(B'|\xi) \, \mathrm{d}G_1(\xi), \tag{19}$$

and tomorrow's participation constraints given poaching and following the mechanism's outcome. That is, whenever $B(\xi) \geq B'^{+} + \xi$, $R_1(B',\xi) \geq B'^{+} + \xi$. Moreover, it must hold that $\Pi_1(R_0), \Pi_1(R_1) \geq \Pi_0$, and $R_0 \geq 0$.

Lemma A.1. *Assuming that $u'(w)$ is bounded below by $\underline{u}'$, the change in $T[\Pi_1]$ is bounded by $-\frac{1}{\underline{u}'}$.*

Proof. Consider two values R and $R' > R$ and their respective strategies. We use all the elements of the strategy at R' but change the wage so that the worker receives a value of R. This implies a strictly lower wage, and provides a lower bound on the expected profit of the firm at R, which is strictly larger than $\hat{\Pi}_1(R)$. Hence $\hat{\Pi}_1$ is strictly decreasing: $\hat{\Pi}_1(R') - \hat{\Pi}_1(R) < -\frac{R'-R}{\underline{u}'}$. □

Lemma A.2. *Assume* Π_1 *is strictly decreasing and* $\Pi_1(0) > \Pi_0$. *Define* $\Pi_1(S) = \Pi_0$. *Then, for all* ξ, $B(\xi) = S$ *is a non-dominated credible strategy.*

Proof. Take a feasible policy, $\theta = \{w, \tilde{\delta}, R_0, R_1, B\}$ with $B(\xi_0) \neq S$ for some $\xi = \xi_0$. First, let's rule out $B(\xi_0) > S$. Credibility imposes that the bid is feasible even when the incumbent looses. This directly imposes $B(\xi_0) \leq S$.

We then show that $B(\xi_0) < S$ is weakly dominated by a strategy that chooses $B(\xi_0) = S$. Construct the alternative policy $\hat{\theta} = \{\hat{w}, \hat{\tilde{\delta}}, \hat{R}_0, \hat{R}_1, \hat{B}\}$ that is identical to θ for all ξ and B', except when $\xi = \xi_0$ in which case $\hat{B}(\xi_0) = S$ and $\hat{R}_1(B', \xi_0) = \max\{B'^+ + \xi_0, R_1(B', \xi_0)\}$. Let's look at the different cases for $B'^+ + \xi_0$ and show that it is feasible and at least as good as θ.

Case 1: when $B'^+ + \xi_0 \leq B(\xi_0)$, the incumbent wins the auction in both θ and $\hat{\theta}$ since $\hat{B}(\xi_0) > B(\xi_0)$. In this case R_1 and $\hat{R}_1$ are the same. The worker and the firm are indifferent between the two strategies.

Case 2: when $B'^+ + \xi_0 \geq \hat{B}(\xi_0)$ the incumbent looses the auction in both θ and $\hat{\theta}$ and get Π_0. The worker moves to the poacher which must deliver at least as much as the bids. The bid is larger in $\hat{theta}$ than under θ, hence he would prefer $\hat{theta}$.

Case 3: Suppose that $B(\xi_0) < B'^+ + \xi_0 \leq S$. Under strategy θ, the incumbent loses and receives Π_0 and the worker receives $\max\{B(\xi_0), \xi_0\}$. Under strategy $\hat{\theta}$, $\hat{B}(\xi_0) = S \geq B'^+ + \xi_0$. The incumbent wins the auction and pays $B'^+ + \xi_0$ to the worker. The worker prefers strategy $\hat{\theta}$, as $B'^+ + \xi_0 \geq \max\{B(\xi_0), \xi_0\}$. The firm also prefers strategy $\hat{\theta}$, as $B'^+ + \xi_0 \leq S$ implies $\Pi_1(B'^+ + \xi_0) \geq \Pi_1(S) \geq \Pi_0$ (since Π_1 is decreasing).

Overall, $\hat{\theta}$ is weakly preferred by the worker, and hence is feasible since θ is feasible. At the same time, it is weakly preferred by the firm. For any feasible strategy with $B(\xi_0) < S$, there is a weakly preferred strategy $B(\xi 0) = S$. We conclude that $B(\xi_0) = S$ is non dominated. □

Lemma A.3. *If* $\Pi_1(R)$ *is* $\gamma(1-\beta)$ *strongly concave (and hence almost everywhere* *differentiable) and weakly decreasing, and $g(v) = -u^{-1}(v)$ is γ-strongly concave, then $\hat{\Pi} = T[\Pi_1]$ is $\gamma(1-\beta)$ strongly concave.*

Proof. Using $v = u(w)$ and $g(v) = -u^{-1}(v)$, we have assumed that g is γ-strongly concave. We note that the solution is γ-strongly concave when $\tilde{\delta}$=1, so we focus our attention on the substantive part of $\tilde{\delta}$=δ. The constraints in the Lagrangian are

$$\mu\Big(v-c-\frac{r}{1+r}W_0+\frac{(1-\lambda_1)(1-\delta)}{1+r}R_0+\frac{(1-\delta)\lambda_1}{1+r}\iint\Big[\mathbf{1}\left\{S\geq B'^{+}+\xi\right\}R_1(B',\xi)$$
$$+\mathbf{1}\left\{S<B'^{+}+\xi\right\}\max\left\{S,\xi\right\}\,\mathrm{d}F(B'|\xi)\,\mathrm{d}G_1(\xi)-R\Big)$$
$$+\frac{(1-\delta)\lambda_1}{1+r}\iint\mu_1(B',\xi)\mathbf{1}\left\{S\geq B'^{+}+\xi\right\}\left(R_1(B',\xi)-B'^{+}-\xi\right)\,\mathrm{d}F(B'|\xi)\,\mathrm{d}G_1(\xi)$$

Where $\mu_1(B',\xi)\geq 0$ is the multiplier for $R_1(B',\xi)-B'^{+}-\xi\geq 0$ and we abstract from the additional simpler bound constraints on R_1 and R_0. We extract the FOCs:

$$g'(v)+\mu=0, \qquad \Pi_1'(R_0)+\mu=0,$$
$$(\mu_1(B',\xi)+\Pi_1'(R_1(B',\xi))+\mu)\,\mathbf{1}\left\{S\geq B'^{+}+\xi\right\}=0.$$

Since R_1 is irrelevant whenever $\mathbf{1}\left\{S\geq B'^{+}+\xi\right\}=0$ we impose that it takes the value S in such case and otherwise we have:

$$\mu_1(B',\xi)+\Pi_1'(R_1(B',\xi))+\mu=0.$$

We assumed that Π_1 is weakly decreasing, which implies that $-\mu$ is also weakly decreasing in R from the F.O.C. on R_0. This tells us that $R_1(B',\xi)$ is at least weakly increasing in R (since it is constant when the constraint binds or $\Pi_1'(R_1(B',\xi))=-\mu$ when it does not, and Π_1 is strictly concave). This implies that increasing R will make the constraint $R_1(B',\xi)-B'^{+}-\xi\geq 0$ more slack, which in turn means that the multiplier $\mu_1(B',\xi)$ will decrease in R. Using finite difference notation $\Delta_R f(R) = f(R')-f(R)$, we get:

$$\begin{aligned}-\Delta_R\mu &= \Delta_R\Pi_1'(R_1(B',\xi))+\underbrace{\Delta_R\mu_1(B',\xi)}_{\leq 0}\\ &\leq \Delta_R\Pi_1'(R_1(B',\xi))\\ &\leq \gamma(1-\beta)\Delta_R R_1(B',\xi)\end{aligned}$$

similary, since Π_1 is $(1-\beta)\gamma$-strongly convave, and $g(\cdot)$ is γ-strongly concave, we have by definition that

$$-\Delta_R\mu=\Pi_1'(\Delta_R R_0)\leq\gamma(1-\beta)\Delta_R R_0$$

and

$$-\Delta_R\mu = \Delta_R g(v) \leq \gamma \Delta_R v$$
$$-\gamma \Delta_R v \leq \Delta_R \mu.$$

We then turn to the PK constraint, we have

$$\Delta_R R = \Delta_R v + \frac{(1-\lambda_1)(1-\delta)}{1+r}\Delta_R R_0 + \frac{(1-\delta)\lambda_1}{1+r}\iint \Big[\mathbf{1}\left\{S \geq B'^{+} + \xi\right\} \Delta_R R_1(B', \xi)\,\mathrm{d}F(B'|\xi)\,\mathrm{d}G_1(\xi),$$

We multiply by $\gamma(1-\beta)$ and use that the different $1/(1+r)$ terms sum to some $p < 1$ together with the derived inequalities to get:

$$p\beta(-\Delta_R\mu) + \gamma\,(1-\beta)\,\Delta_R v \leq \gamma\,(1-\beta)\,\Delta_R R$$
$$p\beta(-\Delta_R\mu) \leq \gamma\,(1-\beta)\,\Delta_R R + (1-\beta)\,\Delta_R\mu$$
$$-\Delta_R\mu\,(\beta p + 1 - \beta) \leq \gamma\,(1-\beta)\,\Delta_R R$$
$$\underbrace{-\Delta_R\mu\beta(p-1)}_{\geq 0} - \Delta_R\mu \leq \gamma\,(1-\beta)\,\Delta_R R$$
$$-\Delta_R\mu \leq \gamma\,(1-\beta)\,\Delta_R R$$

We then go to the envelop condition that gives us

$$\hat{\Pi}(\Delta_R R) = -\Delta_R\mu \leq \gamma\,(1-\beta)\,\Delta_R R$$

which establishes that $\hat{\Pi}(R)$ is $\gamma\,(1-\beta)$-strongly concave. □

Lemma A.4. *The operator T maps $L_\infty(\mathbb{S})$ into $L_\infty(\mathbb{S})$.*

Proof. Proving this only requires showing that the image through the operator is bounded. The expression under the sup operator is directly bounded by

$$T[\Pi_1](R) \leq f - \underline{w} + \frac{1}{1+r}\max\{||\Pi_1||_\infty, \Pi_0\}.$$

and so it is bounded above.

Regarding the lower bound, we need to show that there is a feasible strategy that delivers a bounded value. We consider the strategy that fires the worker while choosing the wage to deliver exactly R: $\tilde{\delta} = 1$ and $u(w) = R + c + \frac{r}{1+r}W_0$. This is feasible since none of the constraints on promises or bids must hold. This provides a lower bound for

$$T[\Pi_1](R) \geq f - u^{-1}(\bar{S} + c + \frac{r}{1+r}W_0) + \frac{1}{1+r}\Pi_0.$$

This implies that $T[\Pi_1]$, the image of Π_1, is bounded, and therefore the image is in $L_\infty(\mathbb{S})$. □

Lemma A.5. *T is a contraction on $L_\infty(\mathbb{S})$.*

Proof. We prove Blackwell's sufficient conditions. We start with monotonicity. Let $\Pi_1 \leq \Pi_1'$ be two bounded functions. Making Π_1 larger relaxes the constraint on the space of strategies since it only appears to impose that promises need to be such that $\Pi(R) \geq \Pi_0$. Therefore, any strategy feasible under Π_1 will be feasible under Π_1'. And any sequence of such strategies leading to the sup will also be feasible. Second, Π_1 enters linearly and positively in the expression under the sup in (18). The sup under Π_1' will then be at least as large as under Π_1, demonstrating monotonicity.

Next, we turn to discounting. Let $\alpha \geq 0$ and $\Pi_1'(R) = \Pi_1(R) + \alpha$. Again, this relaxes the constraints, so any feasible strategy under Π_1 will be feasible under $\Pi_1 + \alpha$. We start with the optimal strategy at $\Pi_1 + \alpha$ (or the sequence of strategies leading to the sup) and collect the terms in the expression under the sup that contain the constant α. This reveals the expression under the sup evaluated with Π_1 without α, at the optimal strategy (or the sequence of strategies) for $\Pi_1 + \alpha$, plus the constant $\frac{1-\tilde{\delta}}{1+r}\alpha$. By definition of the sup, it must be less than $T[\Pi_1]$. Given that $\tilde{\delta} \geq \delta$, this gives us that:

$$T[\Pi_1 + \alpha](R) \leq T[\Pi_1](R) + \frac{1-\delta}{1+r}\alpha.$$

It follows that T is a contraction with modulus $\beta = \frac{1-\delta}{1+r}$. □

It follows that there exists a unique solution Π_1 to the firm problem in $L_\infty(\mathbb{S})$.

Lemma A.6. *The fixed point solution $T(\Pi_1) = \Pi_1$ exists, is unique, continuous, strictly concave, and strictly decreasing.*

Proof. We have shown that the operator is a contraction, and we have shown that it maps strongly concave functions into strongly concave functions. The fixed point is then also strongly concave, and hence strictly concave. The uniqueness of the solution follows from the contraction property. Strictly decreasing follows from Lemma A.1. Continuity follows from concavity. □

Lemma A.7. *We finally show that:*

1. *When the incumbent keeps the worker, the continuation values are given by* $R_1(B', \xi) = \max\{R, B'^{+} + \xi)\}$ *and* $R_0 = R$.
2. *In the interior region, we have* $\Pi_1'(R) = -\frac{1}{u'(w(R))}$, *which implies that the wage next period remains as in the current period whenever* $R_1(B', \xi) = R$.

Proof. We write the Lagrangian for the remaining arguments. Specifically:

$$\begin{aligned}\mathcal{L} &= f - w + \frac{\tilde{\delta}}{1+r}\Pi_0 + \frac{(1-\tilde{\delta})(1-\lambda_1)}{1+r}\Pi_1(R_0) \\ &- \mu\left(u(w) - c - \frac{r}{1+r}W_0 + \frac{(1-\lambda_1)(1-\tilde{\delta})}{1+r}R_0 - R\right) + \\ &\frac{(1-\tilde{\delta})\lambda_1}{1+r}\iint \mathcal{H}(B', \xi, R_1(B', \xi), \mu, \bar{\mu}(B', \xi), \underline{\mu}(B', \xi))\, \mathrm{d}F(B'|\xi)\, \mathrm{d}G_1(\xi)\end{aligned}$$

where $\mathcal{H}$ is function of 6 scalars:

$$\begin{aligned}\mathcal{H}(B', \xi, R_1, \mu, \bar{\mu}, \underline{\mu}) &= \mathbf{1}_{\{S \geq B'^{+}+\xi\}}\Pi_1(R_1) + \mathbf{1}_{\{S < B'^{+}+\xi\}}\Pi_0 - \mu\mathbf{1}_{\{S \geq B'^{+}+\xi\ \}}R_1 \\ &- \mu\mathbf{1}_{\{S < B'^{+}+\xi\}}\max\{S, \xi\} + \mathbf{1}_{\{S \geq B'^{+}+\xi\}}\left(\underline{\mu}\cdot(R_1 - B'^{+} - \xi) + \bar{\mu}\cdot(R_1 - S)\right)\end{aligned}$$

We introduced the two sets of constraints that whenever $B'^{+} + \xi \leq S$ we must impose that $B'^{+} + \xi \leq R_1(B', \xi) \leq S$. The lower bound is captured by multiplier $\underline{\mu}(B', \xi)$ while the upper bound is captured by $\bar{\mu}(B', \xi)$. We scaled both multipliers by $\frac{(1-\tilde{\delta})\lambda_1}{1+r} f(B'|\xi)g_1(\xi)$ to make things readable. Finally, let μ be the Lagrange multiplier of the promise-keeping constraint. We ignore $R_1, R_0 \geq 0$ and verify it holds for the optmial choice. The first-order conditions for w and R_0 are

$$-1 - \mu u'(w) = 0 \text{ and } \Pi_1'(R_0) - \mu = 0.$$

To maximize with respect to $R_1(B', \xi)$ we apply a particular case of the Euler-Lagrange theorem, one that does not require the derivative of the control and states that we can simply take a derivative of the $\mathcal{H}$ function. It is also equivalent to directly checking the small deviations of the $R_1(B', \xi)$ as in a simepl case of a Frechet derivative. We get the following conditions:

$$\mathbf{1}_{\{S \geq B'^{+}+\xi\}}\left(\Pi_1'(R_1(B', \xi)) - \mu + \bar{\mu}_{B',\xi} + \underline{\mu}(B', \xi)\right) = 0,$$

First, whenever the firm loses the auction, then the promised value is not relevant. When the firm wins the auction, if the lower bound of the value is binding, then

$R_1(B', \xi) = B'^{+} + \xi$, if the upper bound is binding, then $R_1(B', \xi) = S$. In the interior region, we get that

$$\Pi_1'(R_1(B', \xi)) = \mu = \Pi_1'(R_0),$$

which by strict concavity gives $R_1(B', \xi) = R_0$. Since we also have $R_0 \leq S$, we get $R_1(B', \xi) = \max\{R_0, B'^{+} + \xi\}$. When $B'^{+} + \xi > S$, we can assume that $R_1(B', \xi) = S$.

The envelope condition gives us $\Pi_1'(R) = \mu = \Pi'(R_0)$ which by the strict concavity of Π_1 tells us that $R_0 = R$. Together with the first-order condition for the wage, it also tells us that $\Pi_1'(R) = -\frac{1}{u'(w(R))}$. □

A.2 Identification Step 3

A.2.1 Distributions of transition wages

Lemma A.8. *We show that the wage after a move for a worker x moving from a firm y to a firm y' has support in $[w(x, y', 0), w(x, y', S(x, y'))]$ with CDF given by:*

$$F_{\text{JJ}}(w|x, y, y') = \frac{\overline{G}_1(S - R(x, y', w))}{\overline{G}_1(S(x, y) - S(x, y'))} \tag{20}$$

Proof. We use the result from the model that the value a worker receives after a move is given by $R' = \max\{S(x, y) - \xi, 0\}$, and the worker collects ξ in addition. This implies that the wage is given by $w(x, y', R')$. We treat the cases $R' = 0$ and $R' > 0$ separately. Because the move has occurred, the distribution of ξ is conditional on $S(x, y') + \xi > S(x, y)$, and therefore the CDF of starting wages is $\overline{G}_1(\xi)/\overline{G}_1(S(x, y) - S(x, y'))$.

We know that $w(x, y', R')$ is monotone in R', and therefore the lowest wage offered will be $w(x, y', 0)$. We have $R' = 0$ iff $\xi > S$. This then implies that there is a mass point at $w(x, y', 0)$

$$F_{\text{JJ}}(w(x, y', 0)|x, y, y') = \mathbb{P}[\xi > S(x, y)] = \overline{G}_1(S(x, y))/\overline{G}_1(S(x, y) - S(x, y')).$$

When $\xi < S$, we have $R' = S(x, y) - \xi$ and for any wage $w \in]w(x, y', 0), w(x, y', S')[$, the CDF following a job-to-job transition is given by

$$\begin{aligned}
F_{\text{JJ}}(w|x, y, y') &= \mathbb{P}[w_{t+1} \leq w | x, y_t = y, y_{t+1} = y', m_t = \text{JJ}] \\
&= \mathbb{P}[w(x, y', R') \leq w | x, y_t = y, y_{t+1} = y', m_t = \text{JJ}] \\
&= \mathbb{P}[w(x, y', S(x, y) - \xi) \leq w | S(x, y') + \xi > S(x, y)] \\
&= \overline{G}_1(S(x, y) - R(x, y', w))/\overline{G}_1(S(x, y) - S(x, y')).
\end{aligned}$$

Therefore, we have for all $w \in [w(x,y',0), w(x,y',S(x,y'))[$ that $F_{\text{JJ}}(w|x,y,y') = \overline{G}_1(S(x,y) - R(x,y',w))/\overline{G}_1(S(x,y) - S(x,y'))$. $\square$

Lemma A.9. *We show that* $F_{\text{UE}}(w|x,y') = \overline{G}_0(-R(x,y',w))/\overline{G}_0(-S(x,y'))$.

$$\begin{aligned} F_{\text{UE}}(w|x,y') &= \mathbb{P}[w_{t+1} \leq w|x, y_{t+1} = y', m_t = \text{UE}] \\ &= \mathbb{P}[w(x,y',-\xi) \leq w|\xi > -S(x,y')] \\ &= \overline{G}_0(-R(x,y',w))/\overline{G}_0(-S(x,y')). \end{aligned}$$

A.2.2 Proof of Lemma 3

We can then prove Lemma 3. We start with the result from Lemma A.8 which gives that

$$F_{\text{JJ}}(w|x,y,y') = \overline{G}_1(S(x,y) - R(x,y',w))/\overline{G}_1(S(x,y) - S(x,y')).$$

Define $z = S(x,y) - R(x,y',w)$ and replace $w = w(x,y',S(x,y) - z)$ to get

$$F_{\text{JJ}}\Big(w(x,y',S(x,y) - z)|x,y,y'\Big) = \overline{G}_1(z)/\overline{G}_1(S(x,y) - S(x,y')),$$

for all $z \in [S(x,y) - S(x,y'), S(x,y)]$.

First, consider within-group JJ-transitions from $y_t = y$ to $y_{t+1} = y$. Such a move requires a draw $0 < \xi \leq S(x,y)$. Under the assumption that $G_1(0) = 1/2$, we therefore get:

$$F_{\text{JJ}}(w(x,y,S - z)|x,y,y) = 2\,\overline{G}_1(z).$$

Hence, $G_1(z)$ is identified on $[0, \overline{S}]$, where $\overline{S} = \max S(x,y)$.

A.2.3 Proof of Lemma 4

We proceed in multiple steps. We start by considering the probability of a move between two firms, which is given by

$$p_{\text{JJ}}(y'|x,y) = \overline{\delta}(x,y)\lambda_1 \frac{v(y')}{V} \overline{G}_1(S(x,y) - S(x,y')),$$

We then consider the following ratio

$$\frac{p_{\text{JJ}}(y'|x,y)}{p_{\text{JJ}}(y''|x,y)} = \frac{v(y')}{v(y'')} \frac{\overline{G}_1(S(x,y) - S(x,y'))}{\overline{G}_1(S(x,y) - S(x,y''))}$$

where the ratio of $\overline{G}_1()$ is identified from Lemma 3 and knowledge of $S(x,y)$ from Lemma 2. The left-hand side of the equation is also known. This shows us that $v(y')/v(y'')$ is known for all pairs of viable matches. Provided sufficient overlap between matches (that is, there is a path between all y' and y'') we identify all ratios. Since $\sum_y v(y) = V$, this means that we have also identified $v(y)/V$ for all y.

Knowing $v(y')/V$, the expression for p_{JJ} identifies $\overline{\delta}(x,y)\lambda_1$. We then move to expressing the probability of separation to unemployment:

$$\mathbb{P}[m_t = \mathrm{EU}|x, y_t = y] = \delta(x,y) + \overline{\delta}(x,y)\lambda_1 \left(\int (1-\phi(x,y')) \frac{v(y')}{V} \, \mathrm{d}y' \right) \overline{G}_1(S(x,y)),$$

where the left hand side if known from Step 1. The right-hand term of the sum is also known since $\phi(x,y)$, $\overline{G}_1$ and $v(y')/V$ are known, and $\overline{\delta}(x,y)\lambda_1$ is known from the previous paragraph. This gives us that $\delta(x,y)$ is identified. We also see that λ_1 is identified by going back to the expression for p_{JJ}.

A.2.4 Proof of Lemma 5

We start by writing down the probability for transitioning from unemployment into being employed in firm y for each worker type x:

$$\mathbb{P}[y_{t+1} = y, m_t = \mathrm{UE}|x] = \lambda_0 \frac{v(y)}{V} \overline{G}_0(-S(x,y))$$

where we know $v(y)/V$ which means that $\lambda_0 \overline{G}_0(-S(x,y))$ is known. We then turn to the distribution of starting wages from unemployment from Lemma A.9 given by:

$$F_{\mathrm{UE}}(w|x,y) = \overline{G}_0(-R(x,y,w))/\overline{G}_0(-S(x,y)).$$

We introduce $\xi = -R(x,y,w)$ and use $R(x,y,w(x,y,-\xi)) = \xi$ to write

$$\forall \xi \in [-S(x,y), 0] \qquad \lambda_0 \overline{G}_0(\xi) = \frac{V}{v(y)} \mathbb{P}[y_{t+1}{=}y, m_t{=}\mathrm{UE}|x] F_{\mathrm{UE}}(w(x,y,-\xi)|x,y),$$

which identifies $\lambda_0 \overline{G}_0(\xi)$ for $\xi \in [-\bar{S}, 0]$.

A.3 Proof of Lemma 6: Identification Step 4

We use equation (2) at $R = S(x,y)$:

$$r\Pi_1(x,y,S(x,y)) = r\Pi_0(y) = (1+r)\left(f(x,y) - w(x,y,S(x,y))\right)$$

where we can directly express the production function in terms of known quantities and $\Pi_0(y)$ in

$$f(x,y) = w(x,y,S(x,y)) + \frac{r}{1+r}\Pi_0(y). \tag{21}$$

The final step is to express the present value of a vacancy where everything is known except for the total number of vacancies V.

$$
\begin{aligned}
rV\Pi_0(y) = \lambda_0 \iint_{-S(x,y)}^{0} \frac{\overline{G}_0(\xi)}{u'(w(x,y,-\xi))} \ell_0(x)\, \mathrm{d}\xi\, \mathrm{d}x \\
+ \lambda_1 \iiint_{-S(x,y)}^{0} \frac{\overline{G}_1(\xi + S(x,y'))}{u'(w(x,y,-\xi))} \overline{\delta}(x,y) \ell_1(x,y')\, \mathrm{d}\xi\, \mathrm{d}x\, \mathrm{d}y'.
\end{aligned} \tag{22}
$$

Hence, we know $\Pi_0(y)$ up to a scale constant V. Finally, we identify the constant V by matching the labor share in the data, i.e. the ratio of average labor expenditure to average revenue $\mathbb{E}[w_{it}]/\mathbb{E}[f(x_i, y_{j(i,t)})]$. We have that

$$
\text{labor share} = \mathbb{E}\left[w_{it}\right]/\mathbb{E}\left[w(x_i, y_{j(i,t)}, S(x_i, y_{j(i,t)})) + \frac{1}{V}\frac{rV}{1+r}\Pi_0(y)\right], \tag{23}
$$

where $V\Pi_0(y)$ is known from equation (22), the labor share is observed directly, and all other objects are also known besides V. This pins down V and, consequently, $\Pi_0(y)$. $f(x,y)$ is then known by equation (21).

We then focus on identifying $\tilde{c} = c(x,y) + b(x)$. $\tilde{c}(x,y)$ is the disutility of labor net of amenities of an (x,y)-match plus the forgone home production. Using the equation for the match surplus we have

$$
\tilde{c}(x,y) = u(\overline{w}(x,y)) + \frac{\overline{\delta}(x,y)\lambda_1}{1+r} \int_{S(x,y)}^{\infty} \overline{G}_1(\xi)\, \mathrm{d}\xi + \left[r + \delta(x,y)\right] S(x,y),
$$

where all elements besides $\tilde{c}(x,y)$ are already identified. Finally, we note that knowing V we also know $v(y)$ since we identified $v(y)/V$ in Lemma 4. With knowledge of $v(y)$ we identify the total mass of jobs in the economy by adding vacant and active jobs $n(y) = \int \ell_1(x,y)\, \mathrm{d}x + v(y))$, which identifies $n(y)$ and concludes the proof of the lemma.

B Online Appendix – not for publication

B.1 Discussion of the Poaching Mechanism of Section 1.3

To fix ideas, consider an ascending auction for the worker's services. Denote the firms' reservation surpluses by $S \equiv S(x, y)$ and $S' \equiv S(x, y')$. As the types are unobserved, reservation surpluses are also private information. Intuitively, an ascending auction should deliver the efficient outcome, and truthfulness should be a dominant strategy for firms.

The ascending auction argument for truthfulness is as follows. First, firms should not make offers above their reservation values S and S', respectively. Second, if the incumbent bids B, then the poacher should bid $B' > B - \xi$, as it knows that the worker will receive an additional ξ when moving. If the poacher bids B', the incumbent should bid $B \geq B' + \xi$, as the worker will need to be compensated for the forgone ξ. Firms continue to make alternating bids until one of them drops out. Therefore, the incumbent retains the worker if $S \geq S' + \xi$ and pays the second price $S' + \xi$. The poacher hires the worker if $S' + \xi > S$ and pays the second price $S - \xi$. When moving to the poacher, the worker receives an additional ξ and thus receives S in total. In the event that $S' = 0$ and $\xi > S$, the worker can choose to leave the incumbent for unemployment and collect ξ. It isconvenient to represent the outcome of this ascending auction as a direct mechanism that resembles the usual second-price sealed bid auction, augmented to account for the mobility shock ξ.

B.2 Additional details for Section 2

The present value for the employed worker The value to a type-x worker employed by a type-y firm at wage w is given by

$$\begin{aligned} W_1(x, y, w) = u(w) - c(x, y) + \frac{\delta(x, y)}{1 + r} W_0(x) + \frac{\overline{\delta}(x, y)\overline{\lambda}_1}{1 + r} \left(W_0(x) + R_0\right) \\ + \frac{\overline{\delta}(x, y)\lambda_1}{1 + r} \iint \Big[\mathbf{1}_{\{B(\xi) \geq B'^{+} + \xi\}} \left(W_0(x) + R_1(B', \xi)\right) \\ + \mathbf{1}_{\{B(\xi) < B'^{+} + \xi\}} \left(W_0(x) + \max\{B(\xi), \xi\}\right) \Big] \, \mathrm{d}F_0(B'|x, \xi) \, \mathrm{d}G_1(\xi), \end{aligned}$$

with $S = S(x,y)$ and $S' = S(x,y')$. Substituting in the equilibrium R_1, R_0, the outcomes of the auction mechanism, and rearranging we obtain:

$$\begin{aligned} rW_1(x,y,w) &= u(w) - c(x,y) + \frac{\delta(x,y)}{1+r}\left(W_0(x) - W_1(x,y,w)\right) \\ &+ \frac{\lambda_1 \overline{\delta}(x,y)}{1+r} \int \Bigg[\int_{W_1(x,y,w)-W_0(x)-S(x,y')}^{S(x,y)-S(x,y')} \left[W_0(x) + S(x,y') + \xi - W_1(x,y,w)\right] g(\xi)\,\mathrm{d}\xi \\ &\quad + \int_{S(x,y)-S(x,y')}^{S(x,y)} \left[S(x,y) + W_0(x) - W_1(x,y,w)\right] g(\xi)\,\mathrm{d}\xi \\ &\quad + \int_{S(x,y)}^{\infty} \left[W_0(x) + \xi - W_1(x,y,w)\right] g(\xi)\,\mathrm{d}\xi \Bigg] \frac{v(y')}{V}\,\mathrm{d}y'. \end{aligned}$$

Let $w(x,y,R)$ be the wage that provides surplus R to the worker. Multiplying by $(1+r)$, subtracting $W_1(x,y,w)$ and $rW_0(x)$ from both sides, and evaluating at $w(x,y,R)$, we obtain the utility flow of the wage equation (7).

And further integrating by part the integrals with respect to ξ, we finally obtain,

$$\begin{aligned} u(w(x,y,R)) &= \frac{r+\delta(x,y)}{1+r}R + \frac{r}{1+r}W_0(x) + c(x,y) \\ &\quad - \frac{\overline{\delta}(x,y)\lambda_1}{1+r}\int_{S(x,y)}^{\infty} \overline{G}_1(\xi)\,\mathrm{d}\xi \\ &\quad - \frac{\overline{\delta}(x,y)\lambda_1}{1+r}\int \left[\int_R^{S(x,y)} \overline{G}_1(R' - S(x,y')^+)\,\mathrm{d}R'\right] \frac{v(y')}{V}\,\mathrm{d}y'. \end{aligned} \tag{24}$$

The minimum wage is $\underline{w}(x,y) = w(x,y,0)$ yielding zero surplus. The maximal wage is $\overline{w}(x,y) = w(x,y,S(x,y))$ yielding maximal surplus $S(x,y)$.

This wage equation is the value of the wage chosen by the firm given types (x,y) and a given worker surplus R (or value $W_0(x) + R$). First, we note that a greater promised surplus R implies a greater wage:

$$\begin{aligned} &(1+r)u'(w(x,y,R))\frac{\partial w(x,y,R)}{\partial R} = \\ &\qquad r + \delta(x,y) + \overline{\delta}(x,y)\lambda_1 \int \overline{G}_1\left[R - S(x,y')^+\right] \frac{v(y')}{V}\,\mathrm{d}y' > 0. \end{aligned} \tag{25}$$

Second, for a fixed R, a greater amenity $-c(x,y)$ requires a lower wage, as in Rosen's (1986) compensating wage differential story. Third, the worker may get an outside offer as a result of holding this job. The wage function depends on y through $c(x,y)$,

$\delta(x, y)$ and $S(x, y)$, and we have

$$(1+r)u'(w(x,y,R))\frac{\partial w(x,y,R)}{\partial S(x,y)} = \\ -\overline{\delta}(x,y)\lambda_1 \int \left(G_1[S(x,y)] - G_1[S(x,y) - S(x,y')^+]\right) \frac{v(y')}{V}\,\mathrm{d}y' < 0. \quad (26)$$

The effect of the current reservation surplus is another compensating differential. A greater surplus decreases the current wage required to achieve the promised value. However, there are two effects that go in opposite directions. By increasing $S(x, y)$ we reduce job-to-job mobility as fewer vacancies can beat the current job. This reduces future payoffs and increases the wage for a given R. But by increasing $S(x, y)$ we also increase the likelihood of future wage raises (in the current job). It happens that the former effect dominates the latter. The intuitions of the basic sequential auction model of Postel-Vinay and Robin (2002) thus carry through to the more general setup, perhaps in a more transparent way.

Surplus equation Substituting expression (8) for the maximum wage along with $R = S(x, y)$ into the wage equation (7) we have

$$u(f(x,y) - r(1+r)^{-1}\Pi_0(y)) = c(x,y) + \frac{r+\delta(x,y)}{1+r}S(x,y) + \frac{r}{1+r}W_0(x) \\ - \frac{\lambda_1\overline{\delta}(x,y)}{1+r}\int \phi(x,y')\left[\int_{S(x,y)-S(x,y')}^{S(x,y)-S(x,y')} [S(x,y') + \xi - S(x,y)]\,g(\xi)\,\mathrm{d}\xi \right. \\ \left. + \int_{S(x,y)-S(x,y')}^{\infty} [\max\{S(x,y)-\xi, 0\} + \xi - S(x,y)]\,g(\xi)\,\mathrm{d}\xi\right]\frac{v(y')}{V}\,\mathrm{d}y'.$$

The first inner integral evaluates to zero. The second inner integral simplifies by noting that when $\xi < S(x, y)$, $\max\{S(x, y) - \xi, 0\} + \xi - S(x, y) = 0$ and when $\xi \geq S(x, y)$, $\max\{S(x, y) - \xi, 0\} + \xi - S(x, y) = \xi - S(x, y)$, so we only need to consider $\xi \geq S(x, y)$, which implies that the inner integral does not depend on y'. The equation defining the surplus then simplifies to equation (9).

Firm profit Substitution in the outcomes of the auction mechanism we can write the firm profit as

$$
\begin{aligned}
[r+\delta(x,y)]\,\Pi_1(x,y,R) &= (1+r)\,[f(x,y)-w(x,y,R)] + \delta(x,y)\Pi_0(y) \\
&+ \overline{\delta}(x,y)\lambda_1 \int \Big[\int_{R-S(x,y')^+}^{S(x,y)-S(x,y')^+} \big(\Pi_1(x,y,S(x,y')^+ + \xi) - \Pi_1(x,y,R)\big)\, \mathrm{d}G_1(\xi) \\
&\qquad + \int_{S(x,y)-S(x,y')^+}^{S(x,y)} (\Pi_0(y) - \Pi_1(x,y,R)) \Big]\, \mathrm{d}G_1(\xi) \frac{v(y')}{V}\, \mathrm{d}y' \Big\}. \qquad (27)
\end{aligned}
$$

Moreover, Theorem 1 establishes that

$$
\frac{\partial \Pi_1(x,y,R)}{\partial R} = -\frac{1}{u'(w(x,y,R))}.
$$

This is a simple consequence of the Envelope Theorem, which can also be obtained by differentiating equation (27) and using equation (25).

If we define the match surplus as the maximal value of R such that $\Pi_1(x,y,R) \geq \Pi_0(y)$, then the profit function being decreasing, $S(x,y)$ is defined by the equality

$$
\Pi_1(x,y,S(x,y)) = \Pi_0(y). \qquad (28)
$$

We can therefore also deduce firm profits from worker wages as

$$
\Pi_1(x,y,R) = \Pi_0(y) + \int_R^{S(x,y)} \frac{\mathrm{d}R'}{u'(w(x,y,R'))}. \qquad (29)
$$

Value of a vacancy Let $J(x,y,R) = \Pi_1(x,y,R) - \Pi_0(y)$ be the gain for the firm of offering a surplus R to the worker. At the equilibrium offers and counter offers the value of a vacancy (3) becomes:

$$
\begin{aligned}
r\Pi_0(y) &= \lambda_0 \frac{L_0}{V} \iint_{-S(x,y)}^{\infty} J(x,y,\max\{-\xi,0\})\, \mathrm{d}G_0(\xi) \frac{\ell_0(x)}{L_0}\, \mathrm{d}x \\
&+ \lambda_1 \frac{L_1}{V} \iiint_{S(x,y')-S(x,y)}^{\infty} J(x,y,\max\{S(x,y')-\xi,0\})\, \mathrm{d}G_1(\xi)(1-\delta(x,y')) \frac{\ell_1(x,y')}{L_1}\, \mathrm{d}x\, \mathrm{d}y'.
\end{aligned}
$$

Breaking the inner integrals into regions above and below the value of ξ at which $\max\{\cdot, 0\}$ becomes 0, we have

$$
\begin{aligned}
r\Pi_0(y) = \lambda_0 \frac{L_0}{V} \int \phi(x,y) \left[\overline{G}_0(0) J(x,y,0) + \int_{-S(x,y)}^{0} J(x,y,0-\xi)\, \mathrm{d}G_0(\xi) \right] \frac{\ell_0(x)}{L_0}\, \mathrm{d}x \\
+ \lambda_1 \frac{L_1}{V} \iint \phi(x,y) \Bigg[\overline{G}_1(S(x,y')) J(x,y,0) \\
+ \int_{S(x,y')-S(x,y)}^{S(x,y')} J(x,y,S(x,y')-\xi)\, \mathrm{d}G_1(\xi) \Bigg] (1-\delta(x,y')) \frac{\ell_1(x,y')}{L_1}\, \mathrm{d}x\, \mathrm{d}y'. \quad (30)
\end{aligned}
$$

Now, consider the two integrals involving ξ in turn. In the first case we have

$$
\begin{aligned}
\int_{-S(x,y)}^{0} J(x,y,-\xi)\, \mathrm{d}G_0(\xi) &= -\left[J(x,y,-\xi)\overline{G}_0(\xi) \right]_{-S(x,y)}^{0} - \int_{-S(x,y)}^{0} \frac{\partial J}{\partial R}(x,y,-\xi)\overline{G}_0(\xi)\, \mathrm{d}\xi \\
&= -J(x,y,0)\overline{G}_0(0) + \int_{-S(x,y)}^{0} \frac{\overline{G}_0(\xi)}{u'(w(x,y,-\xi))}\, \mathrm{d}\xi,
\end{aligned}
$$

where the first equality uses integration by parts and the second equality uses the fact that we defined $J(x,y,S(x,y)) = 0$ and that Lemma (A.7) implies that that $\frac{\partial J}{\partial R}(x,y,R) = -\frac{1}{u'(w(x,y,R))}$. Similarly, in the second case we have

$$
\begin{aligned}
&\int_{S(x,y')-S(x,y)}^{S(x,y')} J(x,y,S(x,y')-\xi)\, \mathrm{d}G_1(\xi) \\
&\quad = -\left[J(x,y,S(x,y')-\xi)\overline{G}_1(\xi) \right]_{S(x,y')-S(x,y)}^{S(x,y')} \\
&\qquad\qquad - \int_{S(x,y')-S(x,y)}^{S(x,y')} \frac{\partial J}{\partial R}(x,y,S(x,y')-\xi)\overline{G}_1(\xi)\, \mathrm{d}\xi \\
&\quad = -J(x,y,0)\overline{G}_1(S(x,y')) + \int_{S(x,y')-S(x,y)}^{S(x,y')} \frac{1}{u'(w(x,y,S(x,y')-\xi))} \overline{G}_1(\xi)\, \mathrm{d}\xi \\
&\quad = -J(x,y,0)\overline{G}_1(S(x,y')) + \int_{-S(x,y)}^{0} \frac{\overline{G}_1(\xi + S(x,y'))}{u'(w(x,y,-\xi))}\, \mathrm{d}\xi.
\end{aligned}
$$

Substituting these two expressions into equation (30) we obtain (11), where, given the bound of the integrals, we can remove $\phi(x,y)$.

B.3 Proof of Lemma 1: Trajectories are Markovian

We treat the different m_t states separately and show that the Markov property holds for each. When the worker is unemployed, we simply write $y_t = 0$ and $w_t = \varnothing$. Our goal is to show

$$\mathbb{P}[w_{t+1}, y_{t+1}, m_{t+1}|x, w_t, y_t, m_t, \Omega_{t-1}] = \mathbb{P}[w_{t+1}, y_{t+1}, m_{t+1}|x, w_t, y_t, m_t]$$

By applying successive conditioning,

$$\begin{aligned}\mathbb{P}[w_{t+1}, y_{t+1}, m_{t+1}|x, w_t, y_t, m_t, \Omega_{t-1}] &= \mathbb{P}[m_{t+1}|w_{t+1}, y_{t+1},\ x, w_t, y_t, m_t, \Omega_{t-1}] \\ &\times \mathbb{P}[w_{t+1}|y_{t+1},\ x, w_t, y_t, m_t, \Omega_{t-1}] \times \mathbb{P}[y_{t+1}|x, w_t, y_t, m_t, \Omega_{t-1}].\end{aligned}$$

We proceed by showing that each of these three probabilities are independent of Ω_{t-1}.

Mobility m_{t+1} It follows from the model that mobility is only a function of the surplus and not of the wage itself.

When an unemployed worker x meets a firm y, the match is formed if and only if $0 \leq -\xi \leq S(x, y)$ (remember that G_0 has negative support):

$$\mathbb{P}\left[m_{t+1} = \mathrm{UE}|x, y_{t+1} = 0, \Omega_t\right] = \lambda_0 \int G_0(-S(x, y))\phi(x, y)\frac{v(y)}{V}\,\mathrm{d}y,$$

and for workers employed in $y_t > 0$ we have (using the short-hand notations $S = S(x, y)$ and $S' - S(x, y')$):

$m_{t+1} = \mathrm{EU}$ if $\xi > \max\{S, S' + \xi\}$:

$$\begin{aligned}&\mathbb{P}\left[m_{t+1} = \mathrm{EU}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] \\ &\qquad = \delta(x, y) + \overline{\delta}(x, y)\lambda_1 \int \left[1 - \phi(x, y')\right]\frac{v(y')}{V}\,\mathrm{d}y' \times \overline{G}_1(S),\end{aligned}$$

$m_{t+1} = \mathrm{JJ}$ if $S' + \xi > \max\{S, \xi\}$ or $S' + \xi > S$ and $S' > 0$:

$$\mathbb{P}\left[m_{t+1} = \mathrm{JJ}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] = \overline{\delta}(x, y)\lambda_1 \int \overline{G}_1(S - S')\phi(x, y')\frac{v(y')}{V}\,\mathrm{d}y',$$

$m_{t+1} = \mathrm{EE}$ if $S \geq S'^{+} + \xi$ or there was no offer:

$$\mathbb{P}\left[m_{t+1} = \mathrm{EE}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] = \overline{\delta}(x, y)\overline{\lambda}_1 + \overline{\delta}(x, y)\lambda_1 \int G_1(S - S'^{+})\frac{v(y')}{V}\,\mathrm{d}y'.$$

Therefore $\mathbb{P}[m_{t+1}|x, w_t, y_t, m_t, w_{t+1}, y_{t+1}, \Omega_{t-1}] = \mathbb{P}[m_{t+1}|x, w_t, y_t, m_t, w_{t+1}, y_{t+1}]$.

Wage w_{t+1} The next step is to examine the law of motion of wages. Whenever unemployed the wage is missing ($w_t = \varnothing$), so we do not need to consider cases $m_t \in \{\mathrm{UU}, \mathrm{EU}\}$ for the law of motion of the wage. We then need to look at UE, EE, and JJ. In each case, we seek an expression for $\mathbb{P}\left[w_{t+1}|x, y_t, w_t, y_{t+1}, m_t, \Omega_{t-1}\right]$.

When $m_t = \mathrm{UE}$, we know that, conditional on moving, the offer is set to deliver a surplus $-\xi$ to the worker, where ξ is a draw from G_0, truncated below by $-S(x, y_{t+1})$. The wage is set through the injective function $w(x, y, R)$ and so:

$$\begin{aligned}\mathbb{P}\left[w_{t+1} \leq w'|x, w_t = w, y_t = 0, y_{t+1} = y, m_t = \mathrm{UE}, \Omega_{t-1}\right] \\ = \mathbb{P}\left[w(x, y, -\xi) \leq w'|\xi \geq -S(x, y)\right] := F_{\mathrm{UE}}(w'|x, y).\end{aligned}$$

Similarly when $m_t = \mathrm{JJ}$, we know that conditional on moving from $y_t = y$ to $y_{t+1} = y'$ the offer is set to deliver a surplus $(S(x, y) - \xi)^+$ to the worker, where ξ is a draw from G_1, truncated below by $S(x, y) - S(x, y')$:

$$\begin{aligned}&\mathbb{P}\left[w_{t+1} \leq w'|x, w_t = w, y_t = y, y_{t+1} = y', m_t = \mathrm{JJ}, \Omega_{t-1}\right] \\ &\quad = \mathbb{P}\left[w(x, y', (S(x, y) - \xi)^+) \leq w' \,|\, \xi \geq S(x, y) - S(x, y')\right] := F_{\mathrm{JJ}}(w'|x, y, y').\end{aligned}$$

Here, we note that in addition it is independent of the previous wage.

We then consider our final case of $m_t = \mathrm{EE}$. In this case, the wage only changes if an outside offer comes in and is above the surplus the worker is getting from their current wage. We know that the surplus the worker receives at wage w from a firm y is equal to $0 \leq R(x, y, w) \leq S(x, y)$. The joint probability of $m_t = \mathrm{EE}$ and $w_{t+1} \leq w' \in [w, w(x, y, S(x, y))]$ for a worker currently employed at a firm $y_t = y$ with a wage $w_t = w$ is the probability of drawing an offer $y_{t+1} = y'$ and a ξ such that $S(x, y')^+ + \xi \leq R(x, y, w')$:

$$\begin{aligned}&\mathbb{P}\left[m_t = \mathrm{EE}, w_{t+1} \leq w'|x, w_t = w, y_t = y, \Omega_{t-1}\right] \\ &\quad = \overline{\delta}(x, y)(1 - \lambda_1) + \overline{\delta}(x, y)\lambda_1 \int G_1\left[R(x, y, w') - S(x, y')^+\right] \frac{v(y')}{V}\, \mathrm{d}y'.\end{aligned}$$

Hence

$$\begin{aligned}&\mathbb{P}\left[w_{t+1} \leq w'|x, w_t, y_t, y_{t+1}, m_t = \mathrm{EE}, \Omega_{t-1}\right] \\ &\quad = \mathbf{1}\left\{w' \geq w\right\} \frac{\mathbb{P}\left[m_t = \mathrm{EE}, w_{t+1} \leq w'|x, w_t = w, y_t = y, \Omega_{t-1}\right]}{\mathbb{P}\left[m_t = \mathrm{EE}|x, y_t\right]} \\ &\quad = \mathbf{1}\left\{w' \geq w\right\} \frac{1 - \lambda_1 \int \overline{G}_1\left[R(x, y, w') - S(x, y')^+\right] \frac{v(y')}{V}\, \mathrm{d}y'}{1 - \lambda_1 \int \overline{G}_1\left[S(x, y) - S(x, y')^+\right] \frac{v(y')}{V}\, \mathrm{d}y'} := F_{\mathrm{EE}}(w'|x, y, w).\end{aligned}$$

This shows that $\mathbb{P}[w_{t+1}|x, w_t, y_t, y_{t+1}, m_t{=}\mathrm{EE}, \Omega_{t-1}] = \mathbb{P}[w_{t+1}|x, w_t, y_t, y_{t+1}, m_t{=}\mathrm{EE}]$.

Firm y_{t+1} Next we turn to y_{t+1}. When $m_t = \mathrm{EU}$, $y_{t+1} = 0$. When $m_t = \mathrm{EE}$, $y_{t+1} = y_t$. When $m_t = \mathrm{UE}$, $\mathbb{P}\left[y_{t+1} = y'|x, m_t = \mathrm{UE}, \Omega_{t-1}\right] \propto \frac{v(y')}{V}\overline{G}_0(-S(x, y'))$. Finally, when $m_t = \text{JJ}$, we get

$$\mathbb{P}\left[y_{t+1} = y'|x, m_t = \text{JJ}, y_t = y, w_t, \Omega_{t-1}\right] \propto \overline{\delta}(x,y)\lambda_1 \frac{v(y')}{V}\phi(x,y')\overline{G}_1(S(x,y)-S(x,y')).$$

This establishes $\mathbb{P}[y_{t+1}|x, w_t, y_t, m_t, \Omega_{t-1}] = \mathbb{P}[y_{t+1}|x, w_t, y_t, m_t]$ and concludes the proof for the Markov property of the model.

B.4 Identification Step 1

We adapt the proof of Bonhomme et al. (2019) to our context of Markovian wages on the job. For simplification, we consider the case with discrete wage outcomes but refer to the original paper for a proof with wages belonging to a continuum.

Throughout the proof, we assume that we have a discretization of the wage where the assumptions hold. This discretization is simply a list of support points w_p for $p \in 1, ..., n_w$. Let also $q \in \{1, ..., n_x\}$ denote the values for worker types x.

Lemma B.1. *We consider 2 firm types y, y' and one middle wage w_2. The distributions*

$$\mathbb{P}[w_1|x, y_1, m_1 = \text{JJ}, m_2 = \text{EE}], \mathbb{P}[w_3|x, w_2, y_2, m_1 = \text{JJ}, m_2 = \text{EE}],$$
$$\mathbb{P}[x, w_2|y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}],$$

are identified from $\mathbb{P}[w_1, w_2, w_3|y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}]$ for all $y_1, y_2 \in \{y, y'\}$ under the assumptions that for any $y_1, y_2 \in \{y, y'\}$,

1. *Wages are Markovian within job spells:*

$$\mathbb{P}[w_3|x, w_2, w_1, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] = \mathbb{P}[w_3|x, w_2, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}].$$

2. *Wages after a move do not depend on wages before the move:*

$$\mathbb{P}[w_2|x, y_1, y_2, w_1, m_1 = \text{JJ}, m_2 = \text{EE}] = \mathbb{P}[w_2|x, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}].$$

3. *The distributions $\mathbb{P}[w_1|x, y_1, m_1 = \text{JJ}, m_2 = \text{EE}]$ and $\mathbb{P}[w_3|x, w_2, y_1, m_1 = \text{JJ}, m_2 = \text{EE}]$ are linearly independent with respect to x for all y_1, w_2 (that is, the CDFs given one x cannot be replicated by the linear combination of the CDFs of w_1 given the other x').*
4. *$d(x, y_1, y_2, w_2) = \mathbb{P}[x, w_2|y_1, y_2, m_1{=}\text{JJ}, m_2{=}\text{EE}] \neq 0$ for all x.*
5. *The following quantity is different for different x's:*

$$\frac{d(x, y, y, w_2)d(x, y', y', w_2)}{d(x, y', y, w_2)d(x, y, y', w_2)}.$$

Proof. We are going to show that given data around a move, we can identify the law of motion for each pair of worker and firm types that employ all types. Throughout y_1 and y_2 can be either y or y'. We can write the following joint density as

$$\begin{aligned}
&\mathbb{P}[w_1, w_2, w_3 | y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] \\
&= \sum_x \mathbb{P}[x | y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] \times \mathbb{P}[w_1, w_2, w_3 | x, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] \\
&= \sum_x \mathbb{P}[x | y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] \times \mathbb{P}[w_1 | x, y_1, y_2, m_1 = \text{JJ}] \\
&\qquad \times \mathbb{P}[w_2 | x, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}] \times \mathbb{P}[w_3 | x, w_2, y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}],
\end{aligned}$$

where we used assumptions i) and ii) to establish the last equality. We then denote the data matrix of the joint density of w_1, w_3 for a fixed value of w_2 and a given y_1, y_2 by $A(y_1, y_2, w_2) \in \mathbb{R}^{n_w \times n_w}$. Hence we have

$$A(y_1, y_2, w_2) = \Big[\mathbb{P}[w_1 \leq w_p, w_2, w_3 \leq w_q | y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}]\Big]_{p,q}.$$

We similarly define the n_w-by-n_x matrices

$$\begin{aligned}
M_1(y_1) &= \Big[\mathbb{P}[w_1 \leq w_p | x = q, y_1, m_1 = \text{JJ}, m_2 = \text{EE}]\Big]_{p,q}, \\
M_{\text{EE}}(y_2, w_2) &= \Big[\mathbb{P}[w_3 \leq w_p | x = q, w_2, y_2, m_1 = \text{JJ}, m_2 = \text{EE}]\Big]_{p,q},
\end{aligned}$$

and the the following diagonal matrix

$$D(y_1, y_2, w_2) = \text{diag}\Big[\mathbb{P}[x = q, w_2 | y_1, y_2, m_1 = \text{JJ}, m_2 = \text{EE}]\Big]_q,$$

where $M_1(y), M_{\text{EE}}(y_2, w_2) \in \mathbb{R}^{n_w \times n_x}$ and $D(y_1, y_2, w_2) \in \mathbb{R}^{n_x \times n_x}$. We can then write the identifying restrictions as

$$A(y_1, y_2, w_2) = M_1(y_1) D(y_1, y_2, w_2) M_{\text{EE}}^\top(y_2, w_2).$$

The next step is to take the singular value decomposition of $A(y, y', w_2) = USV^\top$ where the matrix $S \in \mathbb{R}^{n_x \times n_x}$ by assumption iv) and iii) is diagonal, and $U, V \in \mathbb{R}^{n_w \times n_x}$ are such that $U^\top U{=}V^\top V{=}I_{n_x}$. For all $y_1, y_2 \in \{y, y'\}$, define $B(y_1, y_2, w_2) = S^{-\frac{1}{2}} U^\top A(y_1, y_2, w_2) V S^{-\frac{1}{2}}$ where $B(y_1, y_2, w_2) \in \mathbb{R}^{n_x \times n_x}$ is invertible again by assumption iii) and iv). Note that the different $B(y_1, y_2, w_2)$ matrices for each $y_1, y_2 \in \{y, y'\}$ use the same U, S, V matrices defined as the SVD for $A(y, y', w_2)$.

We first note a property we will use at the end. We have that $D(y, y', w_2) M_{\text{EE}}^\top(y', w_2)$ is full row rank (n_x) by assumption iii) and iv), hence there exist a matrix $\tilde{M} \in \mathbb{R}^{n_w \times n_x}$

such that

$$D(y, y', w_2) M_{\mathrm{EE}}^{\top}(y', w_2) \tilde{M} = I_{n_x}.$$

This implies that

$$\begin{aligned}
UU^{\top} M_1(y) &= UU^{\top} M_1(y) D(y, y', w_2) M_{\mathrm{EE}}^{\top}(y', w_2) \tilde{M} \\
&= UU^{\top} USV^{\top} \tilde{M} \\
&= M_1(y) D(y, y', w_2) M_{\mathrm{EE}}^{\top}(y', w_2) \tilde{M} \\
&= M_1(y).
\end{aligned}$$

We then construct

$$\begin{aligned}
B(y, y, w_2) B(y', y, w_2)^{-1} = {} & S^{-\frac{1}{2}} U^{\top} M_1(y) D(y, y, w_2) M_{\mathrm{EE}}^{\top}(y, w_2) V S^{-\frac{1}{2}} \\
& \times \left(S^{-\frac{1}{2}} U^{\top} M_1(y') D(y', y, w_2) M_{\mathrm{EE}}^{\top}(y, w_2) V S^{-\frac{1}{2}} \right)^{-1} \\
& = Q_1(y) D(y, y, w_2) D(y', y, w_2)^{-1} Q_1(y')^{-1},
\end{aligned}$$

where we used $Q_1(y) = S^{-\frac{1}{2}} U^{\top} M_1(y) \in \mathbb{R}^{n_x \times n_x}$. We note that $Q_1(y')$ is full rank since $A(y', y, w_2)$ has rank n_x from Assumption part iii) and iv). We have then established the following eigenvalue problem:

$$\begin{aligned}
& B(y, y, w_2) B(y', y, w_2)^{-1} B(y', y', w_2) B(y, y', w_2)^{-1} = \\
& \qquad Q_1(y) D(y, y, w_2) D(y', y, w_2)^{-1} D(y', y', w_2) D(y, y', w_2)^{-1} Q_1(y)^{-1}.
\end{aligned}$$

Provided that the eigenvalues are unique, as guaranteed by assumption v), this identifies $Q_1(y)$. We have established that $UU^{\top} M_1(y) = M_1(y)$ and hence we identified $M_1(y) = US^{\frac{1}{2}} Q_1(y)$ up to the scale of the eigenvalue. This scale is pinned down by the fact that the columns of $M_1(y)$ are each a c.d.f which allows using that they equal 1 at w_{n_w}.

With $Q_1(y)$ identified we can use the fact that

$$Q_1^{-1}(y) S^{-\frac{1}{2}} A(y, y', w_2) = D(y, y', w_2) M_{\mathrm{EE}}^{\top}(y', w_2).$$

All objects on the left hand side are known and hence $M_{\mathrm{EE}}(y', w_2)$ is identified up to scale. We can use again, that the columns are CDFs and hence equal 1 at the top.

Once $M_{\mathrm{EE}}(y', w_2)$ and $M_1(y)$ are known we can get $\mathbb{P}[w_2, x | y_1 = y, y_2 = y', m_1 = \mathrm{JJ}, m_2 = \mathrm{EE}]$ from $A(y, y', w_2)$. This gives us the wage conditional on moving, as well as the destination firm for each of the worker types. $\square$

Lemma B.2. *We recover $M_1(y'')$ and $M_{EE}(y'', w)$ for all other y'' and w using the identified $M_{\mathrm{EE}}(y, w_2)$ and $M_1(y)$ from Lemma B.1.*

Proof. We have identified $M_1(y)$ and $M_{EE}(y, w_2)$ for a specific y and w_2. We then note that

$$A(y'', y, w_2)M_{EE}(y, w_2) = M_1(y'')D(y'', y, w_2)M_{\mathrm{EE}}^{\top}(y, w_2)M_{EE}(y, w_2),$$

where $M_{\mathrm{EE}}^{\top}(y, w_2)M_{EE}(y, w_2)$ is know and invertible from Lemma B.1. Hence, the matrix $M_1(y'')D(y'', y, w_2)$ is identified. We finally use the fact that $M_1(y'')$ is a CDF to separate $M_1(y'')$ from $D(y'', y, w_2)$. This identifies $M_1(y'')$ and $D(y'', y, w_2)$ for all y'' with the same labeling as in Lemma B.1. Whenever $D(y'', y, w_2)$ we can't get the corresponding wage density since there are no movers. Though there are no movers only when that particular type never works in the firm.

Next we use the same reasoning for a different w_2. We note that for the y and y' of Lemma B.1 and a w_2' we have:

$$M_1(y)^{\top}A(y, y', w_2') = M_1(y)^{\top}M_1(y)D(y, y', w_2')M_{\mathrm{EE}}^{\top}(y, w_2'),$$

where $M_1(y)^{\top}M_1(y)$ is known and invertible. Hence $D(y, y', w_2')M_{\mathrm{EE}}^{\top}(y, w_2')$ is identified. We finally use the fact that $M_{\mathrm{EE}}^{\top}(y, w_2')$ is a CDF to separate $M_{\mathrm{EE}}^{\top}(y, w_2')$ from $D(y, y', w_2')$. This identifies both. □

Lemma B.3. $\mathbb{P}[y_4, w_4, m_3, y_3, w_3, m_2|m_1 = \mathrm{EE}, y_1, w_1, w_2, x]$ *is identified from* $\mathbb{P}[y_4, w_4, m_3, y_3, w_3, m_2, w_2|m_1 = \mathrm{EE}, y_1, w_1]$ *provided that*

1. $P[w_2, x|m_1 = \mathrm{EE}, y_1, w_1]$ *are linearly independent.*
2. $\mathbb{P}[w_2|x, w_1, y_1, m_1{=}\mathrm{EE}]$ *is known (from Lemma B.1)*

Proof. We first consider the following marginal distribution:

$$\mathbb{P}[w_2|m_1 = \mathrm{EE}, y_1, w_1] = \sum_x \mathbb{P}[w_2|x, m_1 = \mathrm{EE}, y_1, w_1]\mathbb{P}[x|m_1 = \mathrm{EE}, y_1, w_1], \quad (31)$$

where $\mathbb{P}[w_2|x, m_1 = \mathrm{EE}, y_1, w_1]$ is known and the column rank assumption of Lemma B.1 gives that $\mathbb{P}[x|m_1 = \mathrm{EE}, y_1, w_1]$ is identified.

We then note that

$$\begin{aligned}\mathbb{P}[y_4, &w_4, m_3, y_3, w_3, m_2, w_2 | m_1 = \text{EE}, y_1, w_1] \\ &= \sum_x \mathbb{P}[y_4, w_4, m_3, y_3, w_3, m_2 | m_1 = \text{EE}, y_1, w_1, w_2, x] \\ &\qquad \times P[w_2, x | m_1 = \text{EE}, y_1, w_1]\end{aligned}$$

where the left hand side is data and $P[w_2, x|m_1 = \text{EE}, y_1, w_1]$ is known from the previous step and Lemma B.1. The linear independence assumption concludes the proof. □

Corollary 1. $\mathbb{P}[y_4, w_4, m_3, y_3, w_3, m_2|m_1 = \text{EE}, y_1, w_1, w_2, x]$ *identifies the following quantities:*

- $\mathbb{P}[m_t = \text{EU}|x, y_t]$
- $\mathbb{P}[m_t = \text{UE}|x, y_t = 0]$
- $\mathbb{P}[y_{t+1}, w_{t+1}|m_t = \text{UE}, x]$
- $\mathbb{P}[m_t = \text{JJ}|x, y_t]$
- $\mathbb{P}[y_{t+1}, w_{t+1}|m_t = \text{JJ}, x, y_t]$

Proof. The result follows from the Markovian properties of the model. For example:

$$\begin{aligned}\mathbb{P}[y_{t+1}, &w_{t+1}|m_t = \text{UE}, x] \\ &= \mathbb{P}[y_{t+1}, w_{t+1}|m_t = \text{UE}, x, m_{t-2} = \text{EE}, m_{t-1} = \text{EU}, w_{t-2}] \\ &= \mathbb{P}[y_4 = y_{t+1}, w_4 = w_{t+1}, m_3 = \text{UE}, y_3, w_3, m_2|m_1 = EE, y_1, w_1, w_2, x]\end{aligned}$$

□

Corollary 2. *Cross-sectional distributions are identified from using transition probabilities from Corollary 1.*

B.5 Data, sample, and variable construction

We used the raw matched employer-employee data set constructed in Friedrich, Laun, Meghir, and Pistaferri (2019). This links information from three data sources made available by The Institute for Evaluation of Labour Market and Education Policy (IFAU).

The primary data sources for this study are three datasets. The is the Longitudinal Database on Education, Income, and Employment (LOUISE), which provides comprehensive information on demographic and socioeconomic variables for the entire working-age population in Sweden, spanning from 1990 to the present.

The second dataset is the Register-Based Labor Market Statistics (RAMS), which covers employment spells in Sweden starting from 1985 and continuing up to the present time. RAMS includes essential details such as gross annual earnings, the initial and final remunerated month for each employee-firm spell, and unique firm identifiers at the Corporate Registration Number level.

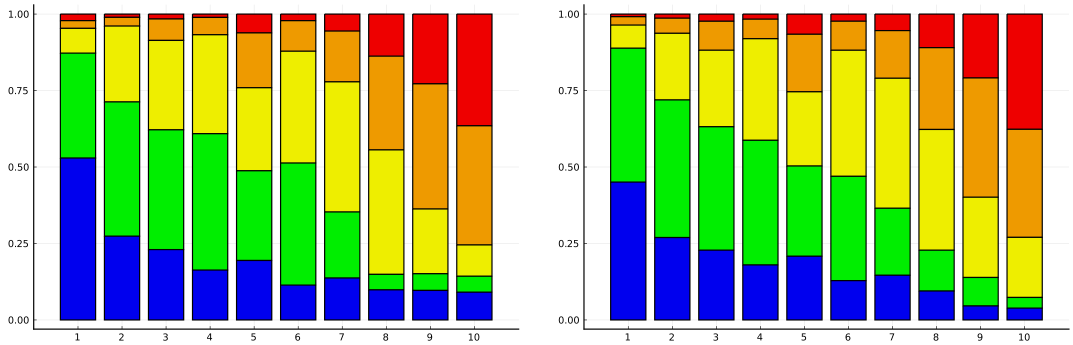
On the firm-related side, RAMS also records information about the industry and the type of legal entity for all firms that employ workers. Finally, we draw from the third data source, the Structural Business Statistics (SBS), which encompasses accounting and balance sheet information for all nonfinancial corporations in Sweden, spanning from 1997 to the present. Of particular interest within SBS is the variable called FORBRUKNINGSVARDE, which provides a measure of value added at both the firm and year level. All monetary variables are adjusted for inflation (detrended with the CPI).

Our analysis is centered on the years 2000 to 2004. The sample we examine comprises all firms classified as either a limited partnership or a limited company, excluding banking and insurance companies. There are two specific restrictions inherited from the original data construction: spells with monthly earnings below 3,416 in 2008 Swedish krona are excluded from the sample, spells that span less than two months of employment (i.e. instances where the start month is the same as the end month) are also excluded from our analysis.

In addition to CPI detrending, we remove year means in the data and we limit it to workers under the age of 50.

B.6 Likelihood for transitions

We estimate the following set of parameters $S(x, y)$ and $\tilde{v}(y) = v(y)/V$, λ_0, λ_1, ρ_0 and ρ_1, that we denote θ by maximizing the following log-likelihood subject to the



(a) Model Stationary

(b) Data (Step 1)

Figure 6: Model stationary distribution and distribution in first step.
Notes: comparing distribution implied by the surplus to distribution estimated with empirical model in Step 1.

moment constraints in the text:

$$
\begin{aligned}
\max_\theta \sum_x \ell_0(x) \Big[& \\
& \sum_{y'} \mathbb{P}[y_{t+1}{=}y', m_t{=}\mathrm{UE}|x] \times \log \lambda_0 \tilde{v}(y') G_0(\rho_0 S(x,y')) \\
& + \mathbb{P}[m_t{=}\mathrm{UU}|x] \times \log \Big(1 - \sum_{y'} \lambda_0 \tilde{v}(y') G_0(\rho_0 S(x,y'))\Big)\Big] \\
+ \sum_{x,y} \ell_1(x,y)\bar{\delta}(x,y) \Big[& \\
& \sum_{y'} \mathbb{P}[y_{t+1}{=}y', m_t{=}\mathrm{JJ}|y_t{=}y, x] \times \log \lambda_1 \tilde{v}(y') G_1(\rho_1 S(x,y') - \rho_1 S(x,y)) \\
& + \mathbb{P}[m_t{=}\mathrm{EE}|y_t{=}y, x] \times \log \Big(1 - \sum_{y'} \lambda_1 \tilde{v}(y') G_1(\rho_1 S(x,y') - \rho_1 S(x,y))\Big)\Big] \\
s.t. \quad m_1(\theta) = m_1, m_2(\theta) = m_2 &
\end{aligned}
$$

where $\mathbb{P}[y_{t+1}{=}y', m_t{=}\mathrm{UE}|x]$, $\mathbb{P}[m_t{=}\mathrm{UU}|x]$, $\mathbb{P}[y_{t+1}{=}y', m_t{=}\mathrm{JJ}|y_t{=}y, x]$, $\mathbb{P}[m_t{=}\mathrm{EE}|y_t{=}y, x]$ and $\ell_1(x, y)$ and $\ell_0(x)$ are known from step 1. The moments $m_1(\theta), m_2(\theta)$ are constructed by simulation, since given θ we can simulate wages and transitions.

References Online Appendix

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70, 2295–2350.

Rosen, S. (1986): "Chapter 12: The theory of equalizing differences," Elsevier, vol. 1 of *Handbook of Labor Economics*, 641–692.